AMERICAN
JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

E. T. BELL
CALIFORNIA INSTITUTE OF TECHNOLOGY

ABRAHAM COHEN
THE JOHNS HOPKINS UNIVERSITY

E. W. CHITTENDEN
UNIVERSITY OF IOWA

F. D. MURNAGHAN
THE JOHNS HOPKINS UNIVERSITY

J. F. RITT
COLUMBIA UNIVERSITY

WITH THE COÖPERATION OF

FRANK MORLEY	MARSTON MORSE	G. C. EVANS
W. A. MANNING	E. P. LANE	AUREL WINTNER
HARRY BATEMAN	ALONZO CHURCH	GABRIEL SZEGÖ
HARRY LEVY	L. R. FORD	R. L. WILDER
J. R. KLINE	OSCAR ZARISKI	R. D. JAMES

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

VOLUME LIX
1937

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND
U. S. A.

matics
FEB 2 - 1937
C-27

AMERICAN JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

E. T. BELL
CALIFORNIA INSTITUTE OF TECHNOLOGY

ABRAHAM COHEN
THE JOHNS HOPKINS UNIVERSITY

E. W. CHITTENDEN
UNIVERSITY OF IOWA

F. D. MURNAGHAN
THE JOHNS HOPKINS UNIVERSITY

J. F. RITT
COLUMBIA UNIVERSITY

WITH THE COÖPERATION OF

FRANK MORLEY	MARSTON MORSE	G. C. EVANS
W. A. MANNING	E. P. LANE	AUREL WINTNER
HARRY BATEMAN	ALONZO CHURCH	GABRIEL SZEGÖ
HARRY LEVY	L. R. FORD	R. L. WILDER
J. R. KLINE	OSCAR ZARISKI	R. D. JAMES

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

Volume LIX, Number 1

JANUARY, 1937

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND
U. S. A.

CONTENTS

PAGE

The cosmical constant and the recession of the nebulae. By Sir Arthur Stanley Eddington, ... 1
The relativistic problem of several bodies. By Tullio Levi-Civita, . 9
Finite ova. By A. R. Poole, ... 23
The integers represented by sets of positive ternary quadratic non-classic forms. By O. K. Sagen, ... 33
Ideal Waring theorem for the polynomial $m(x^3-x)/6-m(x^2-x)/2+x$. By Alvin Sugar, ... 43
Proof of a theorem of Lehmer. By Paul M. Hummel, ... 50
Representation of large integers by cubic polynomials. By Mary Haberzetle, ... 55
Groups of order 64 whose squares generate the four group. By G. A. Miller, ... 57
Necesssary and sufficient conditions for generating certain simple groups by two operators of periods two and three. By Abraham Sinkov, ... 67
Hypergroups. By H. S. Wall, ... 77
Primary Abelian groups and their automorphisms. By Reinhold Baer, 99
Pointwise periodic homeomorphisms. By Deane Montgomery, . 118
Translation groups of three-space. By Deane Montgomery and Leo Zippin, ... 121
Differential properties of abstract transformation groups with abstract parameters. By A. D. Michal and V. Elconin, . . 129
Remarks on systems of ordinary differential equations. By E. R. van Kampen, ... 144
On a symmetrical canonical reduction of the problem of three bodies. By E. R. van Kampen and Aurel Wintner, . . . 153
On the values of the Riemann ζ-function on fixed lines $\sigma > 1$. By Richard Kershner, ... 167
Convolutions of distributions on convex curves and the Riemann zeta function. By E. R. van Kampen and Aurel Wintner, . 175
On the order of the coefficients of a univalent function. By M. S. Robertson, ... 205
Independent sets of postulates for Abelian groups and fields in terms of the inverse operations. By David G. Rabinow, . . 211

The American Journal of Mathematics appears four times yearly.

The subscription price of the Journal for the current volume is $7.50 (foreign postage 50 cents); single numbers $2.00.

A few complete sets of the Journal remain on sale.

Papers intended for publication in the Journal may be sent to any of the Editors.

Editorial communications may be sent to Dr. A. Cohen at The Johns Hopkins University.

Subscriptions for the Journal and all business communications should be sent to The Johns Hopkins Press, Baltimore, Maryland, U. S. A.

Entered as second-class matter at the Baltimore, Maryland, Postoffice, acceptance for mailing at special rate of postage provided for in Section 1103, Act of October 3, 1917, Authorized on July 3, 1918.

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST COMPANY, BALTIMORE, MARYLAND

THE COSMICAL CONSTANT AND THE RECESSION OF THE NEBULAE.[1]

By Sir Arthur Stanley Eddington.

In macroscopic relativity theory energy, momentum and pressure are represented by curvature of space-time; in quantum mechanics they are represented in terms of wave functions. We do not suppose that either the curvature or the waves have an objective existence. It is idle to inquire whether, when we measure energy, we are *really* measuring the bending of space-time in a fifth dimension or the frequency of oscillation of a sub-aetherial fluid. Macroscopic relativity and quantum mechanics are alternative methods of analysing our observational experience. Usually the two methods have different fields of application, and the problems treated by one method would be outside the scope of the other method. But they have a common meeting point; and I shall here consider a problem which can be solved rigorously by both methods. The problem is—to find the state of equilibrium of a radiationless self-contained system of a very large number of particles, positive and negative.

In macroscopic relativity theory this is a familiar problem which was first solved by Einstein. A self-contained static system is called an "Einstein universe." More precisely the system here treated is an Einstein universe with zero pressure; because the absence of radiation implies that the temperature is absolute zero. In quantum mechanics a radiationless steady system is said to be in its "ground state." We have therefore to determine the conditions to be satisfied by a system of N particles treated (a) as an Einstein universe with zero pressure, and (b) as a quantised system in its ground state.

The two answers must agree. But the interesting point is that the relativity solution will be expressed in terms of the gravitational constant κ and the cosmical constant λ, whereas the quantum solution will be expressed in terms of Planck's constant h and other microscopic constants. Thus a comparison of the two solutions will give us a hitherto unrecognized relation between the natural constants.

Surely the determination of this relation is the pre-eminent problem in the unification of macroscopic and microscopic theory—the unification of relativity theory and quantum theory. I shall not be able to give here sufficient

---

[1] A lecture given at the Harvard Tercentenary Conference of Arts and Sciences, Friday afternoon, September 4. Received by the Editors September 3, 1936.

detail for you to check the accuracy of my solution. But I will put before you the principal considerations involved; and if when you come to read the full solution [2] you are dissatisfied with the mathematics, or even (as I have heard whispered) find my treatment obscure, I cannot doubt that you will react in the normal manner of a mathematical physicist who finds an outstanding but obviously soluble problem within his reach—that is to say, you will not rest till you have solved it to your own satisfaction.

It is essential to bear in mind that the curvature and the waves are *alternatives*. We may represent a portion of the energy of a system either by curvature or by waves; but we must not include it twice over in the same formula by representing it both ways. Thus, if we use wave functions to represent part of the energy tensor, these wave functions will occupy a space whose curvature corresponds to the rest of the energy tensor. In particular, if we represent the whole energy tensor by waves, the waves will occupy a flat space. For the solution (b) of the problem enunciated we apply wave representation to the *whole* energy tensor of the system; the waves will accordingly be in flat space.

Thus our first step is to project the spherical Einstein universe into a flat space. For technical reasons it is necessary to use stereographic projection, which has the property that each element of volume remains isotropic after the projection. The uniform distribution of density in spherical space projects into a distribution concentrated towards the centre of projection and fading off to zero at infinity. This resembles the distribution of electron density in an atom. In the atom the density is concentrated towards the centre by the controlling force of the "self-consistent field"; here the controlling force is supplied by a projection factor which has the same effect mathematically. As in an atom, we must analyse the distribution into orthogonal wave functions representing possible steady states of a particle. Then, since the system is in the ground state the particles will occupy the N states of lowest energy.

It will not be necessary to investigate the individual wave functions in detail. We fix attention on two of them, namely the state of lowest energy or K state, and the highest occupied state or limit state. The K state corresponds to uniform distribution in spherical space: it is easy to calculate the wave function of the projected distribution and so determine the pressure and energy (ϵ_1) corresponding to it. The only other result that we require is a very general property of systems in a ground state, namely that when N is large the mean energy per particle is $\frac{3}{5}$ of the limit energy (ϵ_2), so that

[2] Eddington, *Relativity Theory of Protons and Electrons*, Chapter XIV.

(1) $$\bar{\epsilon} = \tfrac{3}{5}\epsilon_2.$$

We already begin to see how things are going to work out. The value of N found from the recession of the spiral nebulae is of order 10^{79}. The limit state will therefore correspond to extremely high quantum number and its energy will be enormously greater than that of the K state. Actually it is found that ϵ_2/ϵ_1 is of order $\sqrt{N}$ or 10^{39}.

We might have begun the problem of unification of relativity and quantum theory by asking ourselves: What is there in common between curvature and periodicity (or frequency) which makes it possible to represent energy sometimes by curvature and sometimes by periodicity? In the simple case of a spherical distribution there is an obvious correspondence of curvature and periodicity—the periodicity being that associated with "going round the world." If there were only one particle in the universe, relativity theory would calculate its energy from the curvature (obtaining the well-known result $\frac{1}{2}\pi Rc^4/\kappa$), and quantum theory would calculate it approximately [3] by treating the circumference of space $2\pi R$ as the wave length of the wave function (obtaining the result $hc/2\pi R$). The two results violently disagree, the ratio being about 10^{120}; but this is not surprising, seeing that there is more than one particle in the actual universe in which h and κ are measured. Even when we divide the curvature result by $N = 10^{79}$, since it corresponds to the energy of N particles each of which should have the energy $hc/2\pi R$ due to the periodicity of its wave function, there remains a huge discrepancy. I think that this discrepancy may have frig tened those who have tried tentatively to find a connection between the two ways of expressing the energy. The reconciliation lies in the fact that the periodicity $2\pi R$ corresponds to the K state of uniform distribution. Only four particles (2 positive and 2 negative) can occupy this state; and the average energy per particle is $\bar{\epsilon}/\epsilon_1$, or about 10^{39}, times larger. The exclusion effect thus completely alters the order of magnitude, and boosts the average energy up to a value which turns out to be intermediate between the energies of constitution m_pc^2, m_ec^2 of protons and electrons.

In elementary quantum theory the system that is being described—the object system—is always treated as an independent addition to the rest of the universe. If the quantum physicist ever remembers that there is a "rest of the universe," he regards it as an ideal background which will not interfere with his problem. When he writes down the ordinary wave equation for say,

[3] The exact calculation by the method of stereographic projection already described yields a result of the same order of magnitude.

two electrons, what does he suppose the other electrons in the universe are doing? Does he suppose them non-existent? If so, it will be difficult for the practical physicist to verify the equation, since he must first destroy the whole universe except two particles. It is clear that the condition that the elementary wave equation shall describe the actual performance of two electrons is, not that the rest of the universe shall be annihilated, but that it shall be arranged in some standard manner. Writers on quantum theory never tell us what standard arrangement is intended; but we can infer it from the assumptions which they commonly make. The rest of the universe must form an impermeable background with no vacant levels to swallow up unexpectedly the two object particles. Nor must any of its particles compete with the object particles for the energy levels in the added system. This means that the rest of the universe must form a system in the ground state with all the levels filled up to a certain limit energy. On top of this comes the added system—*on top*, because, the lower levels being filled, there is no room to add particles except at the top. The limit energy of the background is the threshold energy for the particles of the added system. The particles of the added system have therefore a minimum energy. This is the energy which we recognize as the rest energy or proper mass of the particles.

It must be understood that this arrangement of the rest of the matter of the universe as a system in the ground state, i. e. as an Einstein universe, is not a hypothesis about the actual state of things. (In fact we have reason to believe that the universe is now rather far removed from the Einstein state.) It is the arrangement postulated, when we apply the elementary field-free equations of quantum theory and expect the behaviour indicated by those equations to be exactly fulfilled. Of course, we do not really expect a simple equation to describe what will happen in actual conditions; in actual conditions there are all sorts of perturbing causes—additional protons and electrons, positrons, photons, electromagnetic and gravitational fields—to be taken into account. But when the rest of the universe is not causing perturbations of this kind it must still be accounted for; and it must then form a completely filled series of states below the threshold energy, providing an impermeable foundation for the added system and not otherwise interfering with it. In this way its interference with the object particles is limited to forcing their energies up to the threshold level; that is to say, it is responsible for the proper mass or inertia of the object particles. To use the terminology of relativity theory, the rest of the matter of the universe provides a *pure inertial* field. Any deviation of the rest of the universe from this standard arrangement must be explicitly described as an addition to the fixed background, and in-

cluded as such in the added system. Usually these additions are described as fields. For example, a temperature excitation of the background is described as a *field of radiation*; and exchanges of energy between the object particles and the excited background are described as exchanges between the object particles and a field of radiation incorporated in the added system. In this way we contrive formally to keep the background unchanged, and aloof from what is happening in the added system.

It will be seen that our investigation here links on to Dirac's theory of the positron, provided that certain rather obvious amendments of the latter theory are made. The absence of an electron from a sub-threshold level is represented as the *addition* of a positron to the fixed background. This is virtually Dirac's theory, except that to provide the sub-threshold levels he invented *ad hoc* an infinite quantity of matter undetectable so long as it was uniform. But it is premature to invent unrecognized matter, when there are 10^{79} particles of recognized matter pathetically appealing to the quantum physicist not to neglect them! We emphatically reject this demand for new matter. Other points on which the present theory differs from Dirac's are: (1) The number of sub-threshold levels is not infinite, and we shall presently find the precise number; (2) The levels are filled equally with protons and electrons, so that we avoid Dirac's arbitrary suppression of an infinite negative charge; (3) There is no differential treatment of protons and electrons, so that negatrons are admitted equally with positrons.

We have seen that the rest energy mc^2 of a particle in the added system is the limit energy ϵ_2. Every particle included in the object system means one particle fewer in the "rest of the universe." We can withdraw enough particles to build any system ordinarily studied without making any sensible impression on the huge number N. But cosmical problems form an exception. When we treat the whole universe as our object system, the particles cannot all come from the topmost levels of energy; and the total energy is $N\bar{\epsilon}$ not $N\epsilon_2$, i. e. $\frac{3}{5}Nmc^2$ instead of Nmc^2. The deficit $\frac{2}{5}Nmc^2$ is the amount by which the rest energy of the N particles is less than the sum of the rest energies of the particles taken singly. The ordinary name for such a deficit is gravitational potential energy. Thus $\frac{2}{5}Nmc^2$ is the negative gravitational energy of the matter forming an Einstein universe. Gravitational energy is a kind of negative exclusion effect—due to our having included the *maximum* exclusion energy in the standard mass of the proton or electron. It should be understood that we are not putting forward a new explanation of gravitation; we accepted at the outset the relativity representation of gravitational fields by curvature of space-time. But we have since united relativity and quantum

theory; and we now come across the same thing again, but viewed from the topsy-turvy outlook of elementary quantum theory.

Let us now proceed to the result of the calculation. It is comprised in the formula

$$\sqrt{\tfrac{5}{3}N} = \frac{136}{20}\,\frac{hc}{\kappa(m_p + m_e)^2}. \tag{2}$$

The source of the factor $\frac{5}{3}$ will be evident from what has already been said. The factor $^{136}/_{20}$ is introduced when we take cognizance of the charge and spin of the elementary particles. Perhaps it is simplest to describe it as an averaging factor. Two kinds of particles of unequal mass are present; and for the purposes of the present calculation the equivalent single mass is, not the arithmetic mean, but a more complicated function of the two masses which (for the particular mass-ratio concerned) is $^{2}/_{136}$ of the arithmetic mean.

The cosmical constant λ is connected with the total mass $\frac{1}{2}N(m_p + m_e)$ of the matter in the Einstein universe by the well-known formula [4]

$$\tfrac{1}{2}\pi\lambda^{-\frac{1}{2}} = \tfrac{1}{2}N(m_p + m_e)\kappa/c^2. \tag{3}$$

Thus we can, if we like, eliminate N in (2) and introduce λ in its place. We have then the promised relation between the macroscopic constants κ, λ and the microscopic constants h, m_p, m_e, obtained by solving the same problem in two ways.

Since a knowledge of N is in practice equivalent to a knowledge of λ we look on N as a deputy cosmical constant. To distinguish it, we shall call it the *cosmical number*. Its value can be found from (2) with considerable accuracy, using the observational values of the constants on the right-hand side. We obtain $N = 3{\cdot}145\,.10^{79}$. An instructive way of expressing it is

$$\tfrac{1}{2}N = 135{\cdot}82\,.2^{256}.$$

This suggested to me that the exact value of N might be $2{\cdot}136\,.2^{256}$. I have since been able to show that this is the total number of independent quadruple wave functions of the Dirac type; and an argument from epistemological considerations leaves, I think, no doubt that the principles which we follow in dissecting the sub-stratum of phenomena into elementary particles are such as to yield this total. I feel satisfied therefore that the cosmical number is precisely $2.136.2^{256}$. This number includes both electrons and protons (positrons and negatrons counting negatively in the total). I cannot say definitely

[4] This formula is given by the familiar method of solution, referred to above as method (a).

that the aggregation of some of the particles into atomic nuclei will not introduce any modification. It should however be remembered that we are primarily interested in the number of particles, not in the actual universe, but in the ideal universe postulated in those elementary equations which contain the definitions of m_p, h, etc.

Now let us turn to the extra-galactic nebulae. From (2) and (3) we obtain the cosmical constant and hence by the Friedman-Lemaître theory the limiting speed of recession of the nebulae. We have in fact made two independent determinations, since the value of N has been found in two ways. In the method which I have described at length, we employ the observed values of κ, h, c, m_p. The second method is bolder and employs no observational data, so far as N is concerned. It takes advantage of the fact that *we* lay down the conditions to be fulfilled by a fragment of that maze of interrelatedness which we call the physical universe, to qualify for the status of an independent elementary particle; and in laying down these conditions we (inadvertently) fix the total number of these particles.

From these determinations of the cosmical constant the limiting speed of recession of the spiral nebulae is—

432 kilometres per second per megaparsec.

We must allow some deduction for the gravitational attraction of the nebulae on one another, so that we may expect to observe a recession of 400 km. per sec. per mp., or perhaps rather less. The observed value is usually given as 500 km. per sec. per mp., which is in as close agreement as could be expected.

As a result of this new approach to the problem it is necessary to make two changes in the usual theory of the expanding universe. The mass M_e of an Einstein universe is given in terms of the cosmical constant by the equation $M_e = \frac{1}{2}\pi c^2/\kappa\lambda^{\frac{1}{2}}$. Confining attention to the case of zero pressure, it has been usual to distinguish three possibilities, according as the actual mass M of the universe is greater than, equal to, or less than M_e. In the present theory there is only one case $M = M_e$ to be considered. It happens that this is the choice on which I have plunged in my earlier writings, though admittedly I was guided solely by aesthetic considerations; but now the ambiguity left by relativity theory is definitely settled by quantum theory. Relativity theory gave us no reason to suppose that a collection of particles at zero temperature can form an equilibrium configuration; but quantum theory, appealing to the exclusion principle, assures us that they have an equilibrium configuration, namely the ground state. Thus universes with $M \neq M_e$, which have no

equilibrium configuration, are excluded by quantum theory. In the Friedman-Lemaître theory, it was probably not intended to deny that the universe must in any case have a ground state; but this state was pictured as a kind of "white dwarf" universe in a condition of density outside the scope of the usual cosmical equations. The present calculation shows that, on the contrary, the density of the ground state is of order 10^{-28}.

The second change is in regard to the interpretation of the cosmical constant. It was formerly supposed that λ determined a cosmical curvature $G_{\mu\nu} = \lambda g_{\mu\nu}$, which exists in entirely empty space. It is now clear that $G_{\mu\nu}$ is zero in definitely empty space (i. e. space in which the probability that a particle or photon is present is zero). Definitely empty space is a rather inconvenient abstraction; but we may express the result in a more concrete way by saying that in the space between the galaxies, where the density is much less than the average density of matter in the universe, the curvature is less than $\lambda g_{\mu\nu}$. I cannot enter fully into the reasons for this change of view; but the following hint will perhaps suffice. We had to catechize the quantum physicist, who writes down a wave equation for two or three particles, as to what he had done with the rest of the universe. Similarly when the cosmologist treats the curvature of a vacuum, we have to ask what he has done with the particles removed. In order that we may represent material bodies or vacua as formed by aggregating or dispersing particles (without creation or annihilation of particles) we describe them as modifications of an initial standard probability distribution. In current quantum theory this standard probability distribution (which must not be confused with the standard distribution as an Einstein universe referred to earlier) is known as the *a priori* probability; and the wave functions represent the modifying factor required to transform it into the actual distribution. Since the standard distribution itself is not represented by wave functions, its energy, etc. must be represented by curvature. The cosmical curvature corresponds to this standard or *a priori* probability distribution, and not to a vacuum.

CAMBRIDGE UNIVERSITY,
ENGLAND.

THE RELATIVISTIC PROBLEM OF SEVERAL BODIES.[1]

By TULLIO LEVI-CIVITA.

1. A key stone of classical mechanics is Newton's third law or law of reaction. An immediate consequence of it is that, for every body, the motion of the centre of mass goes on as if this fictitious point were a real particle, acted on all forces *external* to the body. We may therefore disregard, as far as the motion of this point only is concerned, any interior action between the material elements of the body itself.

No analogous statement holds in general relativity. You may still, like in the old mechanics, write down, for every infinitesimal element, the equations of motion corresponding to given physical circumstances. Then, having in view a set of elements, say a natural body, you may form as usual a linear combination of the individual equations, thus introducing the motion of the centre of mass; but in this combination internal forces are no longer eliminated. This indicates that, in general relativity and especially in attacking on astronomical point of view the problem of several bodies, the abstraction of material points (infinitesimal sizes, finite masses) does not offer the same convenience, nor even the same legitimacy, as usual.

Indeed we would meet essential difficulties in searching to pass to the limit too early. The natural way is to start from the general relativistic field-equations, which control the ds^2 and the motion of a *continuous* crowd of material particles.

2. Suppose that, in a given four-dimensional domain of variables, x^0 (timelike), x^1, x^2, x^3 (spacelike) coördinates, there exist a metrics and a flow (vector-field). Mathematically we have then a differential form

$$ds^2 = \sum_0^3{}_{ik} g_{ik} dx^i dx^k; \tag{1}$$

a timelike unit world-vector $\boldsymbol{\lambda}$, whose contravariant and covariant components are respectively

$$\lambda^i = dx^i/ds, \qquad \lambda_i = \sum_0^3{}_k g_{ik}\lambda^k \qquad (i = 0, 1, 2, 3) \tag{2}$$

[1] A lecture given at the Harvard Tercentenary Conference of Arts and Sciences, Friday forenoon, September 4, 1936. Received by the Editors September 3, 1936.

verifying—with the usual notations of the absolute differential calculus—the identities

$$\sum_0^3 {}_i\lambda_i\lambda^i = \sum_0^3 {}_{ik}g_{ik}\lambda^i\lambda^k = 1; \tag{3}$$

and besides a scalar density $\epsilon(x^0, x^1, x^2, x^3)$ fixing at any instant the distribution of energy and therefore also, save for a constant factor, of matter.

Field and motion appear accordingly to involve 15 functions of four independent variables: the ten g, the four λ and ϵ. As they are connected by one algebraic relation (3), we want essentially a system of 14 independent conditions.

If we suppose that only gravitation has to be taken into account, mechanical stresses behave as if the moving matter be incoherent, and the corresponding tensor reduces then to the form

$$T_{ik} = \epsilon\lambda_i\lambda_k \qquad (i, k = 0, 1, 2, 3). \tag{4}$$

With this tensor, Einstein's gravitational equations and their differential consequences afford us the required conditions. Indeed, with well known symbols, the ten gravitational equations

$$G_{ik} - \tfrac{1}{2}Gg_{ik} = -kT_{ik} \qquad (i, k = 0, 1, 2, 3) \tag{I}$$

imply the conservation principles, namely the fact that the vector-divergence of the tensor T_{ik} must vanish, which is expressed by

$$\sum_0^3 {}_k T_{ik}{}^{|k} = 0 \qquad (i = 0, 1, 2, 3).$$

Substituting for T_{ik} their values (4), we get

$$\lambda_i(d\epsilon/ds + \epsilon \operatorname{div} \boldsymbol{\lambda}) + \epsilon p_i = 0, \tag{5}$$

where

$$p_i = \sum_0^3 {}_k\lambda_{ik}\lambda^k \tag{6}$$

are the covariant components of the vector $\boldsymbol{p}$ (geodesic curvature of the flow-line, i. e. trajectory, passing through the point).[2] Equations (5) require firstly that, everywhere matter is present, that is $\epsilon \neq 0$, trajectories are geodesic, in as much as their vector of curvature must vanish. In fact this vector

[2] Compare e. g. my *Absolute Differential Calculus* (Blackie, London and Glasgow, 1927), pp. 135 and 362.

is in any case orthogonal to $\boldsymbol{\lambda}$; therefore multiplying (5) by λ^i and summing, it results, owing to (3),

(III) $$d\epsilon/ds + \epsilon \operatorname{div} \boldsymbol{\lambda} = 0,$$

which is the *equation of continuity* for our gravitational flow. The system (5) reduces then simply to

(II) $$\epsilon p_i = 0 \qquad (i = 0, 1, 2, 3),$$

stating, for $\epsilon \neq 0$, the geodesic property of every trajectory. On the other hand, in empty space ($\epsilon = 0$) trajectories have no physical meaning; they are indeterminate, the counterpart of it being that the equations (II) are automatically satisfied for $\epsilon = 0$.

The four-dimensional equation of continuity (III) gives rise, in a very simple and general manner, to the relativistic correction of the ordinary principle of invariableness of gravitational mass throughout the motion. Indeed (III) is the mathematical expression of the fact that $\int \epsilon dV$ (dV elementary hypervolume) extended to a portion of the fourfold variety (x^0, x^1, x^2, x^3), having the metrical determination (1), is an (integral) invariant of the flow along the geodesic lines. On the other hand the ds^2 is a differential invariant of the same motion, or analytically of the differential system $p_i = 0$ defining the geodesic lines. We have accordingly, along every world-line,

(7) $$\epsilon\, dV/ds = \text{const.}$$

Now the four-dimensional extension dV may be split up in the product

$$dV = \sqrt{-g}\, dx^0 dx^1 dx^2 dx^3 = \sqrt{g/g'}\, dx^0 \sqrt{-g'}\, dx^1 dx^2 dx^3,$$

g and g' indicating respectively the determinants of ds^2 and of its spacial part $(ds^2)_{dx^0=0}$. As

$$\epsilon \sqrt{-g'}\, dx^1 dx^2 dx^3$$

is the energy, or, except for a constant factor, the ordinary gravitational mass dm carried by the moving element, and, with obvious notation,

$$g'/g = g^{00},$$

the preceding equation (7) takes the form

(8) $$\frac{1}{\sqrt{g^{00}}}\, dm \frac{dx^0}{ds} = dm_0 \qquad (dm_0 \text{ constant along the motion})$$

reducing for pseudoeuclidean space to

$$(8') \qquad dm = dm_0 \sqrt{1 - \beta^2}$$

where [3] β means the velocity of the moving particle (x^0 performing the function of time); and the constant dm_0 its ordinary or rest-mass.

In conclusion the mathematical aspect of the most general relativistic problem of the motion of gravitating media consists of 14 partial differential equations between as many unknown functions. As such general media do not exclude the case of separate bodies, ϵ being 0 outside, we have at the same time mathematically sketched even the relativistic problem of several bodies.

3. This rigorous position has been and probably will be successful only in a few cases. Firstly in the celebrated Einstein-Schwarzschild one-centre problem, in which the ds^2, for the space exterior to the centre, may be inferred from (I) through symmetry requirements. Geodesic principle (reduced then to ordinary differential equations) determines afterwards the motion of an *infinitesimal* body in this gravitational field.

Secondly, in the more recent investigations concerning the expanding universe, where some exact solutions have been detected, again in the hypothesis of spherical symmetry (around our immediate surrounding).[4]

Some other exact solutions, especially for static ds^2, have been determined, but their circumstances are not realized, or at least have not hitherto been found realized, in any astronomical question.

4. In order to attack, according to relativity, the usual problems of celestial mechanics, besides the Einstein-Schwarzschild one-centre problem, only approximation's procedures have been tried, and seem to be at present available. They have been initiated by Einstein himself,[5] who integrated, through retarded potentials, the field-equations (I), retaining only linear terms. This has led, among others, to ascertain the existence of gravitational waves.

Almost at the same time the integration by successive approximations of the general equations of moving and gravitating matter has been outlined by *J. Droste* [6] in a thorough but extremely concise paper. The orders of magni-

---

[3] It is not out of place to remark that, for the sake of motion, gravitational mass suffers an alteration (8'), which is exactly inverse of that experienced by inertial mass.

[4] Sources of reference, emphasizing the mathematical point of view, are: G. C. McVittie, *Monthly Notices of R. A. S.*, vol. 93 (1933), pp. 325-339; J. L. Synge, *Proceedings of the National Academy of Sciences*, vol. 20 (1934), pp. 635-640; C. Tolotti, *Rendiconti dell' Accademia dei Lincei*, vol. 21 (1935), pp. 326-331, 488-492, 571-575.

[5] "Näherungsweise Integration der Feldgleichungen der Gravitation," *Sitzungsber. der Preuss. Ak. der Wiss.*, 1916, pp. 688-696.

[6] "The field of moving centres in Einstein's theory of gravitation," *Ak. van Vet. te Amsterdam*, vol. 19 (1926), pp. 447-455.

tude are here previously established with reference to usual conditions of our planetary system; and the process of calculation is shown, for the problem of several, say $n + 1$ bodies, with some useful remarks intended to simplify the laborious task of obtaining, up to the first order, the relativistic corrections in the ordinary differential equations.

This task has been absolved by the late *De Sitter* in his comprehensive researches "On Einstein's theory of gravitation and its astronomical consequences." [7]

In the Newtonian theory, as mentioned at the beginning, the motion of the centre of mass P_h of the h-th body C_h is influenced by the attraction of the other bodies, *not of* C_h *itself*. To get the surrogate of this, even only in the approximate relativistic scheme, is not a simple straightforward matter. De Sitter's survey duly explains the spirit of Einstein's theory, and carefully prepares explicit formulae for eventual astronomical applications. One point however, the contribution of every body to the motion of its material elements, is sketched too hastily. The contributions are not ignored: on the contrary; but the admissions permitting to reduce them to a minimum (of residual constant parameters) are not patiently stated, and even less followed step by step.

Thus it happened that De Sitter was led to the conviction that the self-contributions would only require a slight modification of the constant values of masses, after which no other influences remain in the final equations.

I have recognized after accurate discussion, about which I shall say a few words later on, that the final ignoration of self-contributions is indeed realizable for the problem of the *two* bodies, but not for three or more.

As a matter of fact it may be added that, while the wide interest of De Sitter's work has been soon acknowledged, authoritative voices asked early [8] for a more cogent investigation concerning what, in Newtonian language, would be the residual effects of interior forces on the general motion of a body.

This reserve brought substantially to the diffuse belief [9] that De Sitter's formulae are certainly suitable to govern the events of an *infinitesimal* body, in the field of any others *in given motion*, but not, without caution, to predict the mutual influences of a system of celestial bodies.

---

[7] *Monthly Notices of the R. A. S.*, especially second paper, vol. 67 (December 1916), pp. 155-183.

[8] Compare especially M. Brillouin, "Gravitation Einsteinienne. Statique. Points singuliers. Le point matériel," *Comptes Rendus*, T. 175 (1922), pp. 1008-1012.

[9] This point of view is adopted e. g. in the excellent book of Prof. J. Chazy, *La théorie de la relativité et la mécanique céleste*, Paris, Gauthier-Villars, T. 2 (1930).

5. It is now time to account for the several admissions, which enable firstly the approximate integration of field-equations (I), i. e. the determination of the ds^2; and afterwards the construction of the equations of the motion of the centres of mass $P_0, P_1, \cdots, P_n$ of the bodies, with the aim of getting rid of any other unknown function of x^0 (time) and to reach finally ordinary differential equations as in the classical problem of $n+1$ bodies.

The adopted criteria of simplification are of four kinds and may be summed up as follows:

A_1) Neglect of quantities of order higher than one, in a well defined sense, suggested by the circumstances of our planetary system. Accordingly we shall have to do with velocities of material bodies small in comparison to the velocity of light: order 10^{-4}, as for the earth. For astronomical purposes it will be obviously sufficient to secure the relativistic corrections up to 10^{-8}, and we shall assume this as our need of accuracy: attributing the *first order* to terms having such a small rate to some other standard term in the same equation or relation. We shall generally put, for the ordinary velocity v,

$$v/c = \beta,$$

and say accordingly that β has the order 1/2. In as much the adopted variables are: x^0, which means (nearly) ct (t ordinary time); and x^i $(i > 0)$, which represent (nearly) cartesian coördinates, we may consider for every material particle: dx^i/dx^0 a dimensionless quantity of order 1/2, or $\sim \beta$, $\sim$ meaning "of the same order as . . ."

Furthermore, $f(x^0, x^1, x^2, x^3)$ being a function which varies with x^0 only because of motion of material bodies, $\partial f/\partial x^0$ is of order 1/2 higher than $\partial f/\partial x^i$ $(i > 0)$.

Last but not least comes the behaviour of

$$ds^2 = \sum_0^3 {}_{ik} g_{ik} dx^i dx^k.$$

In empty space the coefficients reduce to the $g^0{}_{ik}$ of the Einstein-Minkowski form $dx^{0^2} - (dx^{1^2} + dx^{2^2} + dx^{3^2})$.

Putting

$$g_{ik} = g^0{}_{ik} - 2\gamma_{ik} \qquad (i, k = 0, 1, 2, 3), \tag{9}$$

all γ_{ik} are to be treated as small quantities of the first order at least; γ_{oi} $(i > 0)$ even of order not beneath 3/2, this arising from the precision with which could be ascertained the isotropy of light propagation through sky, notwithstanding the presence of celestial bodies.

An important complement of these qualitative preliminaries is to avoid superfluous calculations. Aiming to assure the correctness up to the terms of order one in the final equations of motion, it is required, as first remarked by Einstein himself and easy to verify, to procure only the lowest terms of the γ_{ik}, excepting for γ_{00} which is required up to the second order inclusive.

6. Even in the classical mechanics, the problem of several bodies may be reduced to ordinary differential equations only using, besides exact consequences of the mechanical laws, the circumstance (generally, if not always verified) that the dimensions of the bodies are very small in comparison with the mutual distances. The precise assumption, allowing an autonomous treatment of the motion of the centres of mass, is that, D being the largest dimension of any one of the given bodies, and R a lowest bound for mutual distances between points of different bodies, during the motion, not necessarily D/R, but

A_2) $\qquad (D/R)^2$ *is completely negligible.*

As relativity includes ordinary mechanics as limiting case, the same condition A_2) will certainly be required in order to disentangle the motion of the centres of mass. And we shall accordingly associate A_2) to A_1).

There is however an important feature of the ordinary conception, which is lost in applying A_2) to general relativity. In Newtonian mechanics nothing prevents us to suppose the dimension D of the bodies to become smaller and smaller, their mass remaining finite. On the contrary the relativistic determination of the g_{ik} in astronomical problems needs, just in the first approximation, the introduction, for every body C_h, of its own potential, precisely in some point P_h inside C_h, namely

$$\int_{C_h} \frac{\mu' dC'}{r(P', P_h)},$$

where μ' represents the density of the body in its point P', and $r(P', P_h)$ the distance between P' and P_h. Such a value ($> m_h/D$, m_h being the mass of C_h) tends clearly to infinity by lessening the dimensions of C_h and maintaining finite its mass m_h. It is therefore a wise policy to avoid such difficulties, placing ourselves in the circumstances of celestial bodies, for which $(D/R)^2$ may be safely neglected in seeking after first order effects of relativity.

On the other hand it is well to bear in mind that a still greater approximation could not be attainable (as is, on the contrary, the case in ordinary mechanics) by considering bodies of infinitesimal sizes and finite masses: we are then exposed, as just mentioned, to the incongruence of something becoming infinite.

7. The admissions A_1) and A_2) are undoubtedly essential to outline and to perform calculations; they lead especially to the knowledge, within the proposed accuracy, of the ds^2, whose coefficients result as depending on position, velocities and accelerations of all moving elements. It is then possible, applying geodesic principle, to write down, for any one of these elements, the equations of motion. Such equations of course do not yet constitute an ordinary differential system. Truly they concern generic points of continuous media (our $n+1$ bodies) and belong to the much more complicated integro-differential domain.

The further reduction to ordinary equations resembling those of classical mechanics is spontaneously suggested by the last, where, owing to A_2), the motion of centres of mass P_h is rigorously controlled by a system of ordinary differential equations.

The obvious suggestion is to combine the functional equations just arrived at, which define the acceleration of any material element, in order to obtain the acceleration of the various centres of mass P_h $(h = 0, 1, \cdots, n)$. These will consist of Newtonian terms + relativistic corrections. These relativistic corrections, even in the reduced form which follows from A_1), A_2), depend still upon the motion, not only of the P_h's, but in general also of their surrounding bodies C_h. To remove such a priori unknown influences, further hypothesis must be added. Fortunately, as the disturbing bodily influences appear only in the relativistic corrections (order 10^{-8}), it may be anticipated that rough suppositions (verified within one or a few hundredths) will suffice to ensure the required first-order accuracy, as far as 10^{-2} of a first order term is rightly a negligible amount.

The hypothesis, in virtue of which one succeeds in eliminating the functional residua, are two: one, A_3), of kinematical and the other, A_4), of structural contents.

8. A_3) We shall admit that *the motion of every body, say of C_h, is roughly a pure translation.* Roughly means here that, if the vector $\boldsymbol{\beta}_h$ represents the velocity of P_h and $\boldsymbol{\beta}_h + \Delta\boldsymbol{\beta}_h$ the velocity of any other point of the same body C_h, the oscillations are not relevant; more precisely the ratio of the absolute values

$$|\Delta\boldsymbol{\beta}_h| / |\boldsymbol{\beta}_h|$$

never goes beyond a few hundredths. It is certainly so for celestial bodies: for our Earth especially the speed due to rotation at the Equator is scarcely 1/60th of the translational one; not to speak of eventual deformations, whose timerates are in comparison absolutely negligible.

An obvious consequence of A_3) is that, in the same order of approximation, the bodies behave as rigid ones; therefore not only configurations and densities are preserved, but moreover the Newtonian potential of any body C_h will have, in interior points P_h of the same body, invariable values. Accordingly we shall be allowed to treat (in the relativistic corrections) *as independent of* x^0 expressions like

$$\varpi_h = \frac{f}{c^2}\int_{C_h} \frac{\mu' dC'}{r(P', P_h)}\,, \tag{10}$$

where f is the constant of gravitation, μ', as before, the material density.

Except for a constant homogeneity factor (converting the integral into a pure number), ϖ_h is obviously the value of the potential of the h-th body in its centre of mass P_h.

Similarly every C_h will bear almost (always within a few hundredth) unchanged, during the motion, the energy of its own Newtonian distribution, as well—with a far greater accuracy assured by (8′)—its mass m_h. Therefore we are enabled to treat as true constants the dimensionless first order quantities

$$\eta_h = \frac{1}{m_h}\,\frac{f}{c^2}\int_{C_h}\mu(P)\,dC_P \int_{C_h}\frac{\mu(P')}{r(P', P)}\,dC_{P'}, \tag{11}$$

9. The ϖ_h's, just spoken of, though constant with respect to the time x^0, are functions of the interior point P_h, and will thus admit in this point partial derivatives, in general not all zero. Accordingly, in the Lagrangian function belonging to the point P_h, ϖ_h cannot be treated as constant, for, in constructing the equations of motion, the ϖ_h's must be derived just with respect to the coördinates of P_h. Such a complication disappears as soon as the vector $\mathrm{grad}_{P_h}\,\varpi_h$, proportional to Newtonian force, vanishes in P_h. This is obviously the case if the body admits geometrical and material symmetry around one of its points, which is then of course the centre of mass P_h.

But it is not necessary that $\mathrm{grad}_{P_h}\,\varpi_h$ is rigorously zero. It suffices for us that this vector might be neglected in the valuation of relativistic correction. As an essential part of it arises, for every P_h, from the external force due to the other bodies besides C_h, the practical conclusion that in our first order approximation ϖ_h may be treated as a true constant (independent as well of x^0, as of the position of P_h) requires only that:

A_4) *The Newtonian attraction of the body C_h on its centre of mass P_h amounts at most to some hundredth of the attraction exerted, on the same point P_h, by all other bodies of the system.* This will be our fourth and last admission. By the cumulative aid of the four, we are at last enabled to

recognize that the motion of the P_h's finds its reduced mathematical translation in a system of ordinary differential equations of the same total order as the Newtonian problem of the $n+1$ bodies. The ϖ's and η's are to be treated as intrinsic constants of the bodies, well defined in every concrete case of stars or planets. They possess however, as already remarked, the singular character of tending to infinity with the concentration of matter, namely when size decreases, mass remaining finite. This behaviour has no analogue in the ordinary mechanics, but had been noticed in pre-maxwellian and maxwellian electromagnetism and was only recently removed by the new Born's theory.

10. It is not possible, nor would it be suitable to entertain the present audience with the rather long developments leading to an explicit form of the Lagrangian function L_h which furnishes in the usual way the three equations of the motion of the point P_h. I must confine myself to report in synthetical way the final result, to be interpreted as if the surrounding astronomical space be rigorously euclidean; the space-coördinates x^1, x^2, x^3, be cartesian with reference to a whatever galilean frame; the time-coördinate x^0 means ct, as previously stated.

Of course a choice of variables of this kind must be rooted in the very nature of the space-time metrics. I beg you to give me credit that it is really so.

To approach the end I introduce the Newtonian potential, divided by c^2, of the n bodies other than C_h, putting

$$\gamma_h = -\frac{f}{c^2} \sum_0^n{}' {}_\nu \frac{m_\nu}{r(P_h, P_\nu)}, \tag{12}$$

where m_ν denotes the ordinary rest-mass of the body C_ν, and Σ' means that in the sum the term corresponding to the value h of the index ν must be omitted. The coördinates of the point P_h being $x_h{}^i$ $(i = 1, 2, 3)$ and $\dot{x}_h{}^i$ or $\beta_{h/i}$ their derivatives $dx_h{}^i/dx^0$ with respect to x^0, I shall denote by $\boldsymbol{\beta}_h$ the vector-velocity, by

$$\beta_h{}^2 = \sum_1^3 {}_i \dot{x}_h{}^{i2} \tag{13}$$

the square velocity of P_h, and by

$$N_h = \tfrac{1}{2}\beta_h{}^2 + \gamma_h$$

the Lagrangian function of its (absolute) motion in the Newtonian problem of $n+1$ bodies. This function is exactly the classical one, excepted for the

substitution of $x^0 = ct$ to t, which is nothing but a change of unit. It implies division by c^2 of the Newtonian function of force, as shown by the expression (12) of γ_h, which has thus automatically acquired the advantage of being (like β_h) dimensionless.

Coming now to the relativistic correction to N_h, it will be convenient to split it up into two parts D_h and S_h. The first goes up to *Droste* and *De Sitter*[10] (and for that reason I use the letter D) and results from relativistic influences *except self-contributions*, of which the two authors had got rid of with excessive ease.

Professor *M. Brillouin*[11] had happily proposed to name *effacing-principle* (principe d'effacement) the systematic ignoration of self-contribution, ignoration which holds rigorously in Newtonian mechanics, because of equality between action and reaction. In default of this, self-influences must be duly calculated. What, within our approximation, comes out in this way forms the term S_h. The explicit expression of D_h is

$$
\begin{aligned}
D_h = &-\gamma_h{}^2 + \tfrac{1}{2}N_h{}^2 + \gamma_h\beta_h{}^2 \\
&+ \frac{f}{c^2}\sum_0^{n\prime}{}_\nu \frac{m_\nu}{r(P_h, P_\nu)}\,(-\gamma_\nu + \tfrac{3}{2}\beta^2{}_\nu - 4\boldsymbol{\beta}_h \times \boldsymbol{\beta}_\nu) \\
&+ \frac{f}{2c^2}\,\frac{\partial'^2}{\partial' x^{02}}\sum_0^{n\prime}{}_\nu\, m_\nu\, r(P_h, P_\nu),
\end{aligned}
\tag{15}
$$

where $\times$ means scalar product and the dash in the operator $\partial'^2/\partial' x^{02}$ alludes to the circumstance that the derivation does not affect the point P_h, concerning on the contrary all others P_ν $(\nu \neq h)$.

The self-contribution S_h, omitting an additional constant which does not affect the Lagrangian equations of motion, has the form

$$
\begin{aligned}
S_h &= -\tfrac{3}{2}\varpi_h\beta_h{}^2 - \frac{f}{c^2}\sum_0^{n\prime}{}_\nu \frac{m_\nu}{r(P_h, P_\nu)}\,(\eta_\nu + 2\varpi_h) \\
&= -\tfrac{3}{2}\varpi_h\beta_h{}^2 - \frac{f}{c^2}\sum_0^{n\prime}{}_\nu \frac{m_\nu}{r(P_h, P_\nu)}\,\eta_\nu - 2\varpi_h\gamma_h,
\end{aligned}
\tag{16}
$$

η_h and ϖ_h being certain gravitational parameters of the h-th body, which (as already emphasized) play the rôle of (small) constants. By comparing S_h with N_h we see that S_h is built up by exactly the same inertial and gravitational terms as N_h, each affected by a small coefficient ϖ or η. The presence

[10] To be remarked however that *De Sitter* had in view our planetary system, and had accordingly abridged the final expressions of the components of the perturbative forces, treating the masses m_ν $(\nu > 0)$ as infinitesimal in comparison to m_0.

[11] *Loc. cit. ante*, p. 5.

of S_h in L_h has thus barely the effect of altering the inertial and gravitational masses of the bodies. The equality for each body of these two masses is inherent to the Newtonian term N_h, but does no more necessarily hold for the self-terms S_h. Therefore the S_h's will in general produce genuine perturbations of exactly the same order as those already signalled by *De Sitter* and synthetically collected in our D_h.

There is however an important case, the simplest one of the two bodies ($n = 1$), in which a slight alteration of the two masses m_0 and m_1 suffices to reestablish the effacing-principle through elimination of the S_h's. To see this clearly, let me begin by the remark that a Lagrangian function whatever and particularly our

$$L_h = N_h + D_h + S_h$$

may be multiplied by any constant factor without altering the variational equations (equations of motion) arising from it.

We shall employ a factor of the form $1 + \epsilon_h$, ϵ_h being a small constant of the first order. Such a multiplication does not alter, up to the second order, quantities already of second order, like D_h or S_h. We may accordingly assume as Lagrangian function of the motion of P_h instead of L_h,

$$L^*_h = N_h + \epsilon N_h + D_h + S_h$$

with the purpose of opportunely disposing of the first order constants ϵ_h ($h = 0, 1, \cdots, n$).

In the case of two bodies the sums in the expressions (12) and (16) of γ_h and S_h reduce to a single term, bearing, we may say, the index $h + 1$, with the obvious agreement of identifying indices differing by 2 (or a multiple of 2). If furthermore we denote the mutual distance $r(P_h, P_{h+1})$ simply by r, we may then write

$$N_h = \tfrac{1}{2}\beta_h{}^2 + \frac{f}{c^2}\,\frac{m_{h+1}}{r},$$

$$S_h = -\,\varpi_h \tfrac{3}{2}\beta_h{}^2 - \frac{f}{c^2}\,\frac{m_{h+1}}{r}\,(\eta_{h+1} + 2\varpi_h) \qquad (h = 0, 1)$$

from which, adopting $\epsilon_h = 3\varpi_h$ and putting

$$m^*_{h+1} = m_{h+1}(1 - \eta_{h+1} + \varpi_h), \tag{17}$$

which is equivalent to

$$m^*_h = m_h(1 - \eta_h + \varpi_{h+1}), \tag{17'}$$

we get

$$L_h = (1 + 3\varpi_h)N_h + D_h + S_h = \tfrac{1}{2}\beta_h{}^2 + \frac{f}{c^2}\,\frac{m^*_{h+1}}{r} + D_h.$$

In the third term D_h would still appear the old rest-masses m_0, m_1, but it is allowed to replace them by the modified masses m^*_h, since the differences $m^*_h - m_h$ are of the first order and D_h already of the second.

Suppressing asterisks, now useless, because only modified masses subsists, we have finally, for the absolute motion of centres of mass in the problem of the two bodies the Lagrangian functions

$$L_h = N_h + D_h \qquad (h = 0, 1), \tag{18}$$

where $N_h = \frac{1}{2}\beta_h^2 + \gamma_h$ and D_h has the expression (15).

On the contrary, in the problem of three or more bodies, the self-term S_h may not be removed by the simple expedient of such slight mass modifications. At any rate we do not forget the overwhelming circumstance that, even in classical mechanics, the general solution of the problem of several bodies is not known as soon as the bodies are three or more, whilst the two bodies problem had been integrated since Newton himself.

Therefore, though looking only at first order relativistic corrections, the situation is quite different in the general case and in the two bodies problem. For the former only partial investigations may be devised with the aid of some further exact or approximate hypothesis, and so we have meanwhile not attempted to go beyond the precise establishment of differential equations.

For $n = 1$ (two bodies) it follows from the elements of celestial mechanics that only quadratures are needed to obtain first order corrections for the troubled motion, defined by the Lagrangian function (18). It is simply a matter of technical skill to deduce the final formulae in not too tedious a way, thus preparing and discussing possible astronomical tests. I shall briefly sketch and illustrate numerical results on a next occasion.

UNIVERSITY OF ROME.

. . . . . .

In a discussion on the general lines of this paper, which took place October 7 in Princeton, Professor Einstein expressed a preliminary doubt concerning the representation (4) of the T_{ik}. This expression corresponds to incoherent matter, thus disregarding any kind of interior action, normally not at all negligible within matter in bulk.

Professor Einstein remarked that, in order to draw consequences concerning celestial bodies, which behave like concrete aggregates and not as cosmical dust, it seems necessary to introduce in the T_{ik}'s, besides $\epsilon\lambda_i\lambda_k$, which is the relativistic translation of purely kinetic stresses, some further term

concerning material stresses, capable of keeping a body together. The simplest way to do so is to add in the right-hand side of (4) a second term $-pg_{ik}$, which, for $p > 0$, means (speaking the language of classical mechanics) nothing but an isotropic pressure. Such a pressure, resulting from interior actions, has certainly no Newtonian effect on the motions of the centres of mass of the single bodies. It is not self-evident that the influence of pressure is still zero for the relativistic (first order) correction of these motions.

With the feeling that it must be so, I have omitted from the beginning the term $-pg_{ik}$ in the T'_{ik}, taking only later advantage of the circumstance that the bodies to be considered have, almost sensibly, the characteristics of ordinary solids.

Having now extended the analysis of the question, in order to account for p at every step, I just noted that no modification is needed to the conclusions of the preceding investigation, obtaining thus the expected mathematical support to my early dropping of the term in p, which, however, I willingly recognize, does not appear satisfactory from relativistic point of view.

INSTITUTE FOR ADVANCED STUDY,
PRINCETON UNIVERSITY,
November 2, 1936.

FINITE OVA.[1]

By A. R. Poole.

In the following discussion the term finite ovum or simply ovum is used to denote a system consisting of a finite set of distinct elements and a commutative and associative binary rule of combination. The rule of combination is called multiplication and the customary terminology of products and powers is employed. If the ovum has n elements it is said to be of order n. For each ovum of order n there exists a set of $n(n+1)/2$ relations, the multiplication table of the ovum. Two ova of order n are distinct when they are not simply isomorphic.[2] In ovum theory such terms as zero element, identity element, irreducible element, divisor of an element, proper divisor, associate elements, sub-ovum, have the same significance as they have in ring or field theory.

Powers of elements in an ovum. In the chain of elements obtained by taking successive integral powers

$$u_i, u_i^2, u_i^3 \cdots u_i^q \cdots$$

of an element u_i of an ovum, let the p-th power be the first one which gives an element previously occurring in the chain, and suppose

$$u_i^p = u_i^r$$

Let $p - r = s$, so that

$$u_i^{r+s} = u_i^r.$$

There are four possibilities.

$$\begin{array}{lll} (1) & r = 1, & s = 1 \\ (2) & r > 1, & s = 1 \\ (3) & r = 1, & s > 1 \\ (4) & r > 1, & s > 1. \end{array}$$

We shall refer to r as the index and s as the period of the element u_i. If (1) is true u_i is called an idempotent element and evidently all powers of u_i are equal to u_i. If (2) is true we shall call u_i an element of type A. Then

$$u_i, u_i^2, u_i^3 \cdots u_i^r$$

[1] Received September 1, 1936.

[2] Van der Waerden, *Moderne Algebra*, vol. 1, p. 29.

are distinct and

$$u_i^{r+1} = u_i^r.$$

Hence all powers greater than r of u_i are equal to u_i^r. If $u_i^r = u_k$ we shall call u_k the index element of u_i. The element u_k is an idempotent element and is moreover the only idempotent element occurring in the chain of powers of u_i. The elements

$$u_i^t \qquad (t = 1, \cdots, r-1)$$

are elements of type A with indices less than r. If (3) is true we shall call u_i an element of type B. In this case

$$u_i, u_i^2, \cdots, u_i^s$$

are distinct, and

$$u_i^{s+1} = u_i.$$

Evidently

$$u_i^{s+t} = u_i^t \qquad (t \geq 1)$$

and in particular

$$u_i^{2s} = u_i^s$$

so we see that u_i^s is an idempotent element and if

$$u_i^s = u_k$$

we call u_k the period element of u_i. It is easily shown that u_k is the only idempotent element occurring in the chain of powers of u_i and moreover that every element

$$u_i^t, \qquad (t < s)$$

is an element of type B and has the same period element as u_i. If (4) is true we shall call u_i an element of type C. In this case

$$u_i, u_i^2, \cdots, u_i^r, \cdots, u_i^{r+s-1}$$

are distinct and

$$u_i^{r+s} = u_i^r$$

whence for an integer $p \geqq r$ and h any integer

$$u_i^{p+hs} = u_i^p.$$

The elements

$$u_i, u_i^2 \cdots u_i^{r-1}$$

form the unrepeated part of the chain of powers of u_i, the elements

$$u_i^r, \cdots, u_i^{r+s-1}$$

form the repeated part of the chain. Let m be the least integer such that $ms \geqq r$. Then u_i^{ms} is an idempotent element and if

$$u_i^{ms} = u_k$$

then u_k is called the period element of u_i. The element u_k is the only idempotent element occurring in the chain of powers of u_i and it can easily be shown that of the elements occurring in the unrepeated part of the chain those being powers of u_i which are multiples of s are type A elements and the others are of type C, while those other than the period element in the repeated part of the chain are type B elements.

We see that each non-idempotent element in an ovum has one and only one idempotent element in its chain of powers We shall call it the idempotent element of the non-idempotent element.

As an immediate consequence of the preceding we can state

THEOREM 1. *Every ovum contains at least one idempotent element.*

Homomorphisms[3] *in ova containing non-idempotent elements.* In an ovum which contains at least one non-idempotent element, consider the correspondence formed by letting each idempotent element correspond to itself and each non-idempotent element correspond to its idempotent element. Then as such a correspondence is preserved under multiplication we can state

THEOREM 2. *Any ovum which contains at least one non-idempotent element is homomorphic to the sub-ovum formed by its idempotent elements.*

Associate elements in ova. The following theorems on associate elements in ova are easily verified.

THEOREM 3. *In any ovum no two idempotent elements can be associated.*

THEOREM 4. *In any ovum no two non-idempotent elements which have not the same idempotent element can be associated.*

THEOREM 5. *In any ovum a non-idempotent element can not be associated to an idempatent element which is not its idempotent element.*

THEOREM 6. *In any ovum no type A or type C element can be associated*

[3] Van der Waerden, *Moderne Algebra*, p. 32.

to its idempotent element, but every type B element is associated to its idempotent element.

THEOREM 7. *In any ovum no type A or type C element can be associated to an idempotent element or to an element of type B.*

THEOREM 8. *In any ovum two type B elements are associated if and only if they have the same period element.*

THEOREM 9. *In any ovum two type A elements, two type C elements, or an element of type A and an element of type C, can not be associated if they do not have the same index.*

Reduced ova. An ovum in which no pair of elements are associated to one another is called a reduced ovum. From Theorem 6 it follows that a reduced ovum can contain no elements of type B and consequently no elements of type C, and thus contains only elements of type A and idempotent elements We give in detail the proof of the converse of this, namely

THEOREM 10. *Every ovum which contains only idempotent elements and elements of type A is reduced.*

Theorems 3, 4, 7, and 9 show that this result will follow if we prove that in an ovum which contains only idempotent elements and elements of type A, no two type A elements which have the same index element and the same index can be associated. To do this we make use of the following lemmas which are easily verified.

LEMMA 1. *In any ovum if u_i is a type A element with index element u_k, and u_m is another idempotent element such that*

$$u_k u_m = u_m$$

then

$$u_i u_m = u_m.$$

LEMMA 2. *In any ovum, if u_i and u_j are two type A elements with the same index element u_k and the same index, then none of the relations*

$$u_i u_k = u_i$$
$$u_i u_k = u_j$$
$$u_i u_j = u_j$$
$$u_i^2 = u_j$$

are possible.

Returning to the main theorem, let u_i and u_j be two type A elements with index element u_k and index r, in an ovum 0 which contains only idempotent elements and elements of type A. Then

(1) $$u_i u_k = u_k$$

(2) $$u_j u_k = u_k.$$

Assume

(3) $$u_i \sim u_j.$$[4]

Then there exists in 0 an element whose product with u_i is equal to u_j, and by Lemma 2 this element is neither u_i, u_j, or u_k. There must therefore exist in 0 another element say u_l which is such that

(4) $$u_i u_l = u_j.$$

There must also exist in 0 an element whose product with u_j is equal to u_i. From Lemma 2 this element is neither u_i, u_j, or u_k. We show that it cannot be u_l. Assuming

(5) $$u_j u_l = u_i$$

and combining this with (4), we have

(6) $$u_j u_i^2 = u_j$$

(7) $$u_i u_l^2 = u_i.$$

Now

$$u_l^2 = u_l$$

would imply from (6)

$$u_j u_l = u_j$$

in contradiction to (5), so that under assumption (5) u_l^2 cannot equal u_l. From Lemma 2 and (6) and (7) it follows that u_l^2 cannot equal any of u_i, u_j, or u_k. Hence if (5) holds 0 must contain another element u_m such that

(8) $$u_l^2 = u_m$$

(9) $$u_j u_m = u_j$$

(10) $$u_i u_m = u_i.$$

From (8) u_m is either the index element of u_l or is a type A element having the same index element as u_l. Hence for s sufficiently large (i. e. $s \geqq 1$ if u_m is idempotent, $s \geqq t$ where t is the index of u_m if u_m is a type A element)

[4] u_i is associated to u_j.

(11) $$u_l u_m{}^s = u_m{}^s.$$

From (9) and (10) we have

(12) $$u_j u_m{}^s = u_j u_m{}^{s-1} = u_j u_m = u_j,$$

and

(13) $$u_i u_m{}^s = u_i u_m{}^{s-1} = u_i u_m = u_i.$$

Multiplying both sides of (4) by $u_m{}^s$ gives

$$u_i u_l u_m{}^s = u_j u_m{}^s$$

which on employing (11) and (12) gives

$$u_i u_m{}^s = u_j$$

in contradiction to (13). Thus (5) is impossible; so 0 must contain besides u_i, u_j, u_k, and u_l an element u_n which is such that

(14) $$u_j u_n = u_i.$$

From (4) and (14)

$$u_i u_n u_l = u_i$$
$$u_j u_n u_l = u_j.$$

From Lemma 2, $u_n u_l$ can equal neither u_k, u_i, nor u_j; from (4) it cannot equal u_l and from (14) it cannot equal u_n. Hence 0 contains another element u_p which is such that

(15) $$u_n u_l = u_p$$

(16) $$u_i u_p = u_i$$

(17) $$u_j u_p = u_j.$$

We show now that of the three elements u_n, u_p, and u_l no one can be the index element of any other, no pair of them can have the same index element and no one of them has index element u_k. For if u_p is the index element of u_n or if u_n is the index element of u_p or if u_n and u_p both have the same index element, for t sufficiently large ($t \geqq 1$ if u_p is an idempotent element or t greater than or equal to the index of u_p if u_p is a type A element)

(18) $$u_n u_p{}^t = u_p{}^t.$$

Moreover from (16) and (17) we have

(19) $$u_i u_p{}^t = u_i$$

(20) $$u_j u_p{}^t = u_j$$

and from (14)

$$u_j u_n u_p{}^t = u_i u_p{}^t$$

which using (18) and (19) gives

$$u_j u_p{}^t = u_i$$

in contradiction to (20). Similarly u_p cannot be the index element of u_l, u_l cannot be the index element of u_p and u_p and u_l cannot have the same index element. Also if u_n and u_l had the same index element, u_p would be that index element or would have the same index element by (15), in contradiction to what we have just shown. If u_n were the index element of u_l or u_l were the index element of u_n we would have

$$u_n u_l = u_n$$

or

$$u_n u_l = u_l$$

in contradiction to (15).

To show that no one of u_l, u_n or u_p has index element u_k, we first assume that u_p has index element u_k, so that for t greater than or equal to the index of u_p

$$u_p{}^t = u_k$$

whence from (1)

$$u_i u_p{}^t = u_k$$

in contradiction to (19); hence u_p cannot have u_k as its index element. Assume that u_l has index element u_k; then by the above u_n does not have index element u_k, and does not have index element u_p. So u_n is either an idempotent element or there exists in 0 another element u_q which is the index element of u_n. If u_n is an idempotent element, from (14) and Theorem 2, follows

(21) $$u_k u_n = u_k$$

otherwise for u_n a type A element, from (14) and Theorem 2

(22) $$u_k u_q = u_k.$$

From (15) for any integer s

$$u_n{}^s u_l{}^s = u_p{}^s$$

and for s sufficiently large

(23) $$u_n u_k = u_p{}^s$$

if u_n is an idempotent element, otherwise

(24) $$u_q u_k = u_p{}^s.$$

For u_n an idempotent element combining (21) and (23) gives

(25) $$u_p^s = u_k$$

and for u_n a type A element combining (22) and (24) we get (25). But (25) indicates that u_k is the index element of u_p which we have proved impossible. Hence u_l and similarly u_n cannot have u_k as its index element. We have now shown that assumption (3) implies that 0 contains besides the three elements u_i, u_j, and u_k, at least three more elements u_l, u_n, and u_p, and that these six elements satisfy the seven relations

(1) $$u_i u_k = u_k$$
(2) $$u_j u_k = u_k$$
(4) $$u_i u_l = u_j$$
(14) $$u_j u_n = u_i$$
(15) $$u_l u_n = u_p$$
(16) $$u_i u_p = u_i$$
(17) $$u_j u_p = u_j.$$

Moreover of the three elements u_l, u_n, and u_p, no one can be the index element of any other, no pair of them can have the same index element, and no one of them can have index element u_k. Now let u_L denote the index element of u_l if u_l is a type A element, and let u_L denote u_l itself if u_l is idempotent. Let u_N and u_P have similar significance with respect to u_n and u_p. Then from (15) and Theorem 2 follows

(26) $$u_L u_N = u_P.$$

Multiplying both sides of (26) by u_L gives

(27) $$u_P u_L = u_P$$

which by Lemma 1 gives

(28) $$u_P u_l = u_P.$$

From (16) for any integer t

$$u_p{}^t u_i = u_i$$

which for t sufficiently large gives

(29) $$u_P u_i = u_i.$$

Similarly from (17)

(30) $$u_P u_j = u_j.$$

From (4) multiplying both sides by u_P

$$u_i u_l u_p = u_j u_p$$

which on employing (28) and (30) yields

$$u_i u_p = u_j .$$

in contradiction to (29). Thus assumption (3) leads to a contradiction so that we conclude that in an ovum which contains only idempotent elements and elements of type A no type A elements having the same index element and the same index can be associated. Thus the proof of Theorem 10 is now complete.

Further results for reduced ova which are easily verified are:

THEOREM 11. *A reduced ovum must contain a zero element.*

THEOREM 12. *If a reduced ovum has an identity element, that element is an irreducible element.*

COROLLARY 1. *A reduced ovum O containing an identity element has a reduced sub ovum of order $n-1$, consisting of all the elements of O except the identity.*

COROLLARY 2. *From a reduced ovum O of order n mark set $(u_1, u_2 \cdots u_n)$, we can form a reduced ovum of order $n+1$ containing an identity element, by adjoining to the mark set of O an element u_{n+1} and to the multiplication table of O the relations*

$$u^2_{n+1} = u_{n+1}$$
$$u_{n+1} u_i = u_i u_{n+1} = u_i . \qquad (i = 1, 2, 3 \cdots n).$$

THEOREM 13. *A reduced ovum has at least one irreducible element which is not an identity element.*

COROLLARY 1. *Every reduced ovum of order n has at least one reduced sub ovum of order $n-1$.*

This corollary shows that from all possible distinct ova of order $n-1$ we can obtain all possible distinct ova of order n by adjoining to the ova of order $n-1$ another idempotent or type A element making multiplication of this element with itself and with the original elements commutative and associative, and examining the ova thus formed to see which are simply isomorphic to one another.

Ova formed from groups. We consider now those ova which contain no

type A or type C elements but at least one element of type B. For convenience we call these ova of type 2.

If G is a finite abelian group of order $n > 1$ and i is its identity element every element a of G has the property

$$a^n = i$$

whence

$$a^{n+1} = a$$

and so G is an ovum of type 2 with only one idempotent element the identity. Conversely an ovum O of type 2 containing only one idempotent element is a group. More generally we can state

THEOREM 14. *Every ovum O of type 2 is either a group, or consists of sub ova which have no element in common and each of which is a group. Each of these groups consists of an idempotent element and of all the type B elements which have this idempotent element for period element.*

In a group every element divides every other element so that every element is associated to every other element. From Theorems 3 and 6 it follows that an ovum in which every element is associated to every other element can have only one idempotent element and type B elements and is therefore a group. Hence we have

THEOREM 15. *A sufficient condition that an ovum be a finite group is that every element be associated to every other element.*

It must be noted that the condition given in this theorem for finite abelian groups is not as strong as that given by Van Der Waerden in his postulates for groups.[5] His postulate 5 not only demands that every element be associated to every other element, but that every element divide itself.

From Theorem 15 we know that all type 2 ova of order n can be obtained by compounding groups of order $\leq n$, only those combinations of groups being taken the sum of whose orders is equal to n. The commutative and associative laws must be satisfied and the ova must be tested for distinctness. Of interest in the question of forming type 2 ova from groups is

THEOREM 16. *From a finite abelian group of order $n - 1$ we can obtain two and only two ova of order n, by the adjunction of an idempotent element.*

MINNEAPOLIS, MINN.

[5] Van der Waerden, *Moderne Algebra*, p. 19, (5).

THE INTEGERS REPRESENTED BY SETS OF POSITIVE TERNARY QUADRATIC NON-CLASSIC FORMS.[1]

By O. K. SAGEN.

Introduction. A. A. Albert [2] has shown for the classic case that all integers which are not represented by a *set* of positive ternary quadratic forms lie in certain arithmetic progressions. In this paper analogous results are obtained for the non-classic forms by methods similar to Albert's. The problem is solved by determining under what conditions there exist solutions of a Diophantine equation of certain form.

1. *Proper representation by sets.* Consider the positive ternary quadratic non-classic form

$$(1) \qquad f(x,y,z) = ax^2 + by^2 + cz^2 + ryz + szx + txy$$
$$a > 0,\ b > 0,\ c > 0,\ r, s, t \text{ not all even},\ 4bc - r^2 > 0.$$

The determinant of the form $2f(x,y,z)$ is

$$2d = 2(4abc + rst - ar^2 - bs^2 - ct^2).$$

Since the determinant $2d$ is an invariant of the form under unitary transformation we shall define the *set* $\Sigma(d)$ as the set of all forms (1) having the same invariant d. The set $\Sigma(d)$ is said to *represent* an integer m, if there exist integers x, y, z and a form f of invariant d such that $m = f(x,y,z)$. If the g. c. d. of x, y, z is unity, then the representation is called *proper*.

2. *Necessary and sufficient conditions for proper representation.* Suppose that an integer a is properly represented by a set $\Sigma(d)$. It is well known [3] that if an integer a is properly represented by some form, there exists an equivalent form f having a as the coefficient of x^2. Since classic forms are non-equivalent to non-classic forms and since equivalent forms represent the same integers, there exists a positive non-classic form with coefficients a, b, c, r, s, t such that

$$(3) \quad d + ar^2 + bs^2 + ct^2 - rst = 4abc,\ 4bc - r^2 > 0 \ (r, s, t \text{ not all even}).$$

Conversely, if (3) holds then the integer a is properly represented by $\Sigma(d)$.

---

[1] Received January 22, 1934; revised October 15, 1936.

[2] A. A. Albert, "The integers represented by sets of ternary quadratic forms," *American Journal of Mathematics*, vol. 55 (1933), pp. 274-292.

[3] L. E. Dickson, *Studies in the theory of numbers*, 1930. University of Chicago Press, p. 12.

LEMMA. *An integer a is properly represented by a set $\Sigma(d)$, if and only if there exist integers b, c, r, s, t such that* (3) *holds.*

Suppose that (3) is true for given integers a and d. The binary quadratic form $\phi(y, z) = by^2 + ryz + cz^2$ has discriminant $-\Delta = r^2 - 4bc < 0$ and is a positive form, since $b > 0$. By (3), evidently

$$\phi(s, -t) = bs^2 + ct^2 - rst = 4abc - ar^2 - d = a\Delta - d. \tag{4}$$

Let $s_1 = \text{g. c. d. } (s, t)$. Then, on writing $s = s_1 s_0$, $t = s_1 t_0$, $b_1 = \phi(s_0, -t_0)$, we have

$$\phi(s, -t) = s_1^2 b_1. \tag{5}$$

Since ϕ properly represents the positive integer b_1, there exists an equivalent binary ϕ_1 of discriminant $-\Delta$ such that

$$\phi_1(y_1, z_1) = b_1 y_1^2 + r_1 y_1 z_1 + c_1 z_1^2.$$

Therefore there exist integers which satisfy

$$\Delta = 4b_1 c_1 - r_1^2. \tag{6}$$

But, by (4), $b_1 s_1^2 = a\Delta - d$, and consequently (6) implies the existence of integers b_1, c_1, r_1, s_1 such that

$$d + ar_1^2 = b_1(4ac_1 - s_1^2) > 0. \tag{7}$$

By (6) and (4), $r^2 \equiv -\Delta \equiv r_1^2 \pmod 4$; so that, since $s_1 = \text{g. c. d. } (s, t)$, not both r_1 and s_1 are even. Evidently (7) implies (3); so by the lemma we obtain the following theorem.[4]

THEOREM 1. *A positive integer a is represented properly by a set $\Sigma(d)$, if and only if there exist positive integers b, c, r, s, with r, s not both even, such that*

$$d + ar^2 + bs^2 = 4abc. \tag{8}$$

COROLLARY 1. *If equation* (8) *holds with s odd, then r is of arbitrary parity.*

COROLLARY 2. *In equation* (8) *r, s, c may be replaced respectively by $2bv + r$, $2au + s$, $au^2 + bv^2 + c + rv + su$ where u, v are arbitrary integers.*

When (8) is true, evidently

$$d + a(r+s)^2 + (b+a)s^2 + c(2a)^2 - (r+s)s(2a) = 4a(b+a)c. \tag{9}$$

[4] This theorem and its proof are valid for sets of classic forms and from them essential simplifications of Albert's proofs are obtainable.

Since (9) is an equation of the same form as (3), the proof of Theorem 1 shows that integers b_1, c_1, r_1, s_1 exist, such that (7) is true and $r_1^2 \equiv (r+s)^2 \pmod 4$. But since s is odd, we have $r_1 \equiv r+1 \pmod 2$, whence Corollary 1 follows.

Writing $r_1 = 2bv + r$, $s_1 = 2au + s$, $c_1 = au^2 + bv^2 + c + rv + su$, we note that

$$d + ar_1^2 + bs_1^2 = 4abc_1$$

reduces to (8); which verifies Corollary 2.

3. *Necessary and sufficient condition for representation.* Suppose that the positive integers a and d are expressed in their unique forms $a = 4^l\lambda^2 M$, $d = 4^g\Delta$, where λ is odd, M is without square factor, and Δ is not divisible by 4.

THEOREM 2. *A positive integer a is represented by a set $\Sigma(d)$ of non-classic forms, if, and only if, there exist positive integers b, c, r, s such that*

(10) $\Delta + Mr^2 + bs^2 = 4Mbc$, *where s is odd whenever $l < g$, while, if $l \geqq g$, at least one of r and s is odd.*

If a is represented by $\Sigma(d)$, there exists a form f and integers ξ, η, ζ such that $f(\xi, \eta, \zeta) = L^2M$. Let $\theta = \text{g. c. d. } (\xi, \eta, \zeta)$, $\xi = \theta\xi_1$, $\eta = \theta\eta_1$, $\zeta = \theta\zeta_1$. Then $4^l\lambda^2 M = \theta^2 f(\xi_1, \eta_1, \zeta_1)$. Evidently θ is a divisor of $2^l\lambda$, and on writing $2^l\lambda = 2^h\lambda_0\theta$, we note that $4^h\lambda_0^2 M$ is properly represented by $\Sigma(d)$. By Theorem 1,

(11) $$4^g\Delta + 4^h\lambda_0^2 Mr^2 + bs^2 = 4 \cdot 4^h\lambda_0^2 Mbc.$$

If s is odd, on setting $r_1 = 2^h\lambda_0 r$ we have

(12) $$4^g\Delta + Mr_1^2 + bs^2 \equiv 0 \pmod{4Mb}.$$

Otherwise, let $s = 2^\sigma s_1$, where s_1 is odd and $\sigma > 0$. We then obtain

$$4^g\Delta + 4^h\lambda_0^2 Mr^2 + 4^\sigma bs_1^2 \equiv 0 \bmod (4 \cdot 4^h Mb).$$

If $\sigma > h$ then, either $g = h$, whence

(13) $$\Delta + M(\lambda_0 r)^2 + b(2^{\sigma-h}s_1)^2 \equiv 0 \pmod{4Mb},$$

or $g > h$, so that $\sigma = h$, and

(14) $$4^{g-h}\Delta + M(\lambda_0 r)^2 + bs_1^2 \equiv 0 \pmod{4Mb}.$$

If $\sigma \leqq h$, then $g \geqq \sigma$ and, writing $r_2 = 2^{h-\sigma}\lambda_0 r$, we have

(15) $$4^{g-\sigma}\Delta + Mr_2^2 + bs_1^2 \equiv 0 \pmod{4Mb}.$$

But, if $4^k\Delta + Mr_0^2 + bs_0^2 \equiv 0 \pmod{4Mb}$, we may take $r_0 = 2\rho$ in view of Corollary 1, Theorem 1. Hence $b = 4\beta$ and

$$4^{k-1}\Delta + M\rho^2 + \beta s_0^2 \equiv 0 \pmod{4M\beta}.$$

Therefore, (12), (13), (14), (15) imply the existence of integers b^1, r^1, s^1, c^1, such that

(16) $\Delta + Mr^{1^2} + b^1 s^{1^2} = 4Mb^1c^1$, where s^1 is even in case $g = h$, while otherwise s^1 is odd. Hence, if $g > l$, (16) holds with s^1 odd; while for $g \leqq l$, (16) holds with at least one of r^1 and s^1 odd.

Conversely, if the conditions (16) hold for s^1 odd, we may write $r = 2^g r^1$, $b = 4^g b^1$; therefore

$$4^g\Delta + Mr^2 + bs^{1^2} = 4Mbc^1.$$

Thus by Theorem 1, M is properly represented by $\Sigma(d)$. But when s^1 is even, $g = h \leqq l$, and r^1 is odd. Setting $s = 2^g s^1$, we have

$$4^g\Delta + 4^h Mr^{1^2} + b^1 s^2 = 4^{h+1} Mb^1c^1.$$

Therefore $4^h M$ is properly represented by $\Sigma(d)$. By definition then, there exists a form f_1 for s^1 odd such that $f_1(1, 0, 0) = M$, while for s^1 even, a form f_2 exists such that $f_2(1, 0, 0) = 4^h M$. But $f_1(2^l\lambda, 0, 0) = 4^l\lambda^2 M$ and $f_2(2^{l-h}\lambda, 0, 0) = 4^l\lambda^2 M$, so a is therefore represented by $\Sigma(d)$.

4. *Reduction of the general case to a problem for integers without square factors.*

LEMMA 1. *Let Ka be any integer without square factor with K odd, and let d be any integer. There exist solutions b, c, r, s of*

$$K^2 d + Kar^2 + bs^2 = 4Kabc$$

if and only if there exist solutions b^1, c^1, r^1, s^1 of

$$Kd + ar^{1^2} + b^1 s^{1^2} = 4ab^1c^1.$$

Furthermore, s and s^1 are of the same parity.

LEMMA 2. *Let p be an odd prime not dividing a such that either the Legendre symbol $(-ad/p) = +1$, or d is divisible by p. There exist solutions of*

$$p^2 d + ar^2 + bs^2 = 4abc$$

if and only if s, s^1 are of the same parity and

$$d + ar^{1^2} + b^1 s^{1^2} = 4ab^1c^1$$

is solvable.

LEMMA 3. *Let p be an odd prime such $(ad/p) = -1$. Then*

$$p^2 d + ar^2 + bs^2 = 4abc$$

is satisfied if and only if either

$$pd + par^{1^2} + b^1 s^{1^2} = 4pab^1c^1$$

or

$$d + ar^{1^2} + b^1 s^{1^2} = 4ab^1c^1$$

hold with s^1 of the same parity as s.

For the proof of Lemma 1, let the g. c. d. $(b, K) = \beta$, then $K = \beta K_0$, $b = \beta b_1$, $s = K_0 s^1$. By Corollary 2 of Theorem 1, r may be taken divisible by K_0. Hence

$$K^2 d + Kar^2 + bs^2 \equiv 0 \pmod{4KK_0ab},$$

and, on writing $K_0 b_1 = b^1$, we have

$$Kd + ar^2 + b^1 s^{1^2} \equiv 0 \pmod{4ab^1}. \tag{17}$$

Conversely, if (17) holds, let the g. c. d. $(b^1, K) = K_0$ and $K = K_0 K_1$, $b^1 = K_0 b_1$. By Corollary 2, Theorem 1, r may be chosen divisible by K_1 so that (17) holds modulo $4ab^1K$. Thus, setting $s = K_0 s^1$, $b = b^1 K$, and multiplying (17) by K, we obtain

$$K^2 d + Kar^2 + bs^2 \equiv 0 \pmod{4Kab}.$$

Evidently $s \equiv s^1 \pmod 2$, and, hence, Lemma 1 follows.

We prove Lemmas 2 and 3 simultaneously. In case b is not divisible by p, we may, by Corollary 2, Theorem 1, take $r = pr_1$ and $s = ps_1$ so that

$$d + ar_1^2 + bs_1^2 \equiv 0 \pmod{4ab}. \tag{18}$$

But if $b = pb_1$, then $r = pr_1$, and

$$pd + par_1^2 + b_1 s^2 \equiv 0 \pmod{4pab_1}. \tag{19}$$

If $b_1 = pb_2$, then

$$d + ar_1^2 + b_2 s^2 \equiv 0 \pmod{4ab_2}. \tag{20}$$

If b_1 is not divisible by p, then in view of Corollary 2, Theorem 1, r_1 may be chosen so that $d + ar_1^2 \equiv 0 \pmod p$, whenever either p divides d or $(-ad/p) = +1$. Then (19) holds modulo $4p^2ab_1$; therefore,

$$d + ar_1^2 + bs_1^2 \equiv 0 \pmod{4ab}. \tag{21}$$

Conversely, if any of (18), (20), (21) hold, then on multiplying by p^2 we obtain

$$p^2 d + ar^2 + b^1 s^2 \equiv 0 \pmod{4ab^1}.$$

If (19) holds, then, on multiplying by p, we have

$$p^2d + ar^2 + bs^2 \equiv 0 \pmod{4ab}.$$

Evidently $s \equiv s_1 \pmod 2$; hence the lemmas are verified.

Now express Δ of Theorem 2 in its unique form $\Delta = \gamma^2 D$ where γ is odd and D is without square factor. Let $\mu =$ g. c. d. (M, γ), $M = \mu M_1$, $\gamma = \mu\gamma_1$. By Lemma 1, condition (10) holds if and only if it also holds for Δ and M replaced by ${\gamma_1}^2\mu D$ and M_1 respectively. Now let the g. c. d. $(\mu, D) = \mu_0$, $\mu = \mu_0\mu_1$, $D = \mu_0 D_1$, and let P be the product of all primes p dividing $\gamma_1\mu_0$ such that $(-M_1\mu_1D_1/p) = -1$. Then, by Lemma 2, condition (10) holds if and only if it also holds with Δ, M replaced by $P^2\mu_1D_1$, M_1 respectively. Finally, by Lemma 3, (10) holds if and only if it holds either for the pair μ_1D_1, M_1 or for $P\mu_1D_1$, PM_1.

THEOREM 3. *Let $a = 4^l\lambda^2 M$, $d = 4^g\gamma^2 D$,* where *λ, γ are odd, M and D each without square factor. Further suppose $\mu_0 =$ g. c. d. (M, D, γ), $\mu_0\mu_1 =$ g. c. d. (M, γ), $M = \mu_0\mu_1 M_1$, $D = \mu_0 D_1$, $\gamma = \mu_0\mu_1\gamma_1$, and let P be the product of all primes p dividing $\mu_0\gamma_1$ for which $(-M_1\mu_1D_1/p) = -1$. Then a is represented by $\Sigma(d)$ if and only if there exist integers b, c, r, s such that either*

$$\mu_1D_1 + M_1r^2 + bs^2 = 4M_1bc \quad \textit{or} \quad P\mu_1D_1 + PM_1r^2 + bs^2 = 4PM_1bc$$

with s odd, if $l < g$, and with r, s not both even, if $l \geqq g$.

5. *A sufficient condition.*

THEOREM 4. *If $\mu_1 > 1$ or if D_1 has a factor prime to M_1, then a is represented by $\Sigma(d)$ where M_1 and D_1 are defined as in Theorem 3.*

For the proof, let $\Delta = 2^i q\theta$, $N = 2^j\mu$ each be an integer without square factor, such that q is prime to N, and $\mu\theta$ is odd.

First, suppose q is odd and define

$$p_n = [Bq]n + [B^{q-1}(w+1) - 1], \tag{23}$$

such that $B = 8\mu\theta$ and $(w/q) = -(-1/q)$. Since B and q are relatively prime, $B^{q-1} \equiv 1 \pmod q$, by Fermat's theorem. Consequently $p_n \equiv w \not\equiv 0 \pmod q$, $p_n \equiv -1 \pmod B$, and the bracketed expressions in (23) are relatively prime. By Dirichlet's theorem [5] on primes in arithmetic progression v_n represents

[5] P. G. L. Dirichlet, *Abhandlungen der Königlichen Akademie der Wissenschaften*, Berlin, 1837, pp. 108-110.

an infinitude of primes. If, for $n = m$, p_m is an odd prime, write $p_m = p$. We can now compute the Legendre symbol $(-N\Delta/p) = (-2^{i+j}q\theta\mu/p)$. First, $(2/p) = +1$ and $(-1/p) = -1$, since $p \equiv -1 \pmod 8$. Also $(\mu/p) = (\theta/p) = +1$, since $p \equiv -1 \pmod{\theta\mu}$. But $p \equiv w \pmod q$ and hence

$$(q/p) = (-1)^{(p-1)/2\ (q-1)/2}(w/q) = -1.$$

Therefore, $(-N\Delta/p) = +1$, and $(-Q\Delta/p) = +1$, where $NQ \equiv 1 \pmod p$. Consequently, there exists an odd integer r such that $r^2 \equiv -Q\Delta \pmod p$. As a result

$$\Delta + Nr^2 \equiv 0 \pmod p. \tag{24}$$

Writing $s = 1$, $c = 2\theta[qm + B^{q-2}(w+1)]$, we have $p = 4Mc - s^2$.

Now, suppose $q = 2$. Then $\Delta = 2\theta$ and $N = \mu$; hence we write

$$p_n = 8\mu\theta n + 4\mu\theta - 1.$$

By Dirichlet's theorem p_n is an odd prime p, by proper choice of n. Evidently $p \equiv 3 \pmod 8$ and $p \equiv -1 \pmod{\mu\theta}$, and, therefore, $(-1/p) = (2/p) = -1$, $(\mu\theta/p) = +1$. Thus the Legendre symbol $(-N\Delta/p) = +1$, so there exists an odd integer r such that

$$\Delta + Nr^2 \equiv 0 \pmod p. \tag{25}$$

But $p = 4Mc - s^2$, where $c = \theta(2n+1)$ and $s = 1$.

Now if $\mu_1 > 1$ in Theorem 3, or if there exists some factor of D_1 which is prime to M_1, then (24) and (25) imply (22) is satisfied with s odd. Therefore a is represented by $\Sigma(d)$.

6. *Necessary and sufficient conditions for proper representation.*

THEOREM 5. *Suppose AD is without square factor. There exist integers b, c, r, s, with s odd, such that*

$$D + ADr^2 + bs^2 = 4ADbc \tag{26}$$

if and only if there exists an odd integer $q > 1$ which divides D and is such that $(-A/q) = -1$.

Suppose (26) holds and let $q = \text{g. c. d. } (s, D)$, $s = qs_0$, $D = qD_0$, $m = 4AD_0c - qs_0^2$. Evidently m is prime to D_0 and hence, by (26),

$$1 + Ar^2 \equiv 0 \pmod m. \tag{27}$$

Therefore, if $m > 1$, the Jacobi symbol $(-A/m) = +1$. Writing $A = 2^i\alpha$,

where α is odd and i is either 0 or 1, we observe that $(2/m) = (2/q)$ when $i = 1$, since $m \equiv -q \pmod 8$. Also $(-1/m) = -(-1/q)$ and

$$(\alpha/m) = (-1)^{(\alpha-1)/2\,(m-1)/2}\,(m/\alpha)$$
$$= (-1)^{(\alpha-1)/2\,(q+1)/2\,+(\alpha-1)/2+(\alpha-1)/2\,(q-1)/2}.\,(\alpha/q) = +(\alpha/q),$$

since $m \equiv -qs_0^2 \pmod{4\alpha}$. Consequently,

$$(-A/q) = -(-A/m) = -1.$$

If $m = 1$, we have $4AD_0c = qs_0^2 + 1$ and, therefore, $(-q/\alpha) = +1$, since $(qs_0)^2 \equiv -q \pmod\alpha$. But $(-1/q) = -1$, since $q \equiv -1 \pmod 4$, and hence,

$$(-\alpha/q) = (-1)^{1+(\alpha-1)/2\,(q-1)/2\,+(\alpha-1)/2}(-q/\alpha) = -1.$$

If $i = 1$, evidently $q \equiv -1 \pmod 8$ and $(2/q) = 1$. Therefore, again $(-A/q) = -1$.

Conversely, if $D = qD_0$, q odd, $(-A/q) = -1$, we form the expression

$$p_n = 8AD_0n - q. \tag{28}$$

By Dirichlet's theorem, for some value of n, p_n is a prime p. But, since $p \equiv -q \pmod{8A}$, we have $(-1/p) = -(-1/q)$, $(2/p) = (2/q)$ and $(\alpha/p) = (\alpha/q)$. Therefore, the Legendre symbol

$$(-AD_0^2/p) = -(-A/q) = +1,$$

and so there exists an integer r, such that

$$D_0 + AD_0r^2 \equiv 0 \pmod p, \qquad D + ADr^2 \equiv 0 \pmod{qp}. \tag{29}$$

Setting $c = 2n$, $s = q$, evidently $qp = 4ADc - s^2$, and hence (29) implies (26).

THEOREM 6. *If AD is without square factor, $D = 2\delta$, δ odd and $A \equiv 1 \pmod 4$ or $A \equiv 3 \pmod 8$, then there exist integers b, c, r, s such that r is odd, s even, and*

$$D + ADr^2 + bs^2 = 4ADbc. \tag{30}$$

For the proof when $A \equiv 3 \pmod 8$, let

$$p_n = ADn + A\delta - 2.$$

Evidently AD is relatively prime to $A\delta - 2$; hence, by Dirichlet's theorem, there is some value of n for which p_n is a prime p. Since $p \equiv -2 \pmod A$, we find

$$(-A/p) = (-1)^{(p-1)/2+(p-1)/2\,(A-1)/2}(-2/A) = +1.$$

Therefore, there exists an odd integer r, such that

$$1 + Ar^2 \equiv 0 \pmod{p}.$$

But, since $A \equiv 3 \pmod 8$,

$$D + ADr^2 \equiv 0 \pmod{8p}. \tag{31}$$

Now, with $c = 2n + 1$, $s = 4$, evidently $p = 4ADc - s^2$, and therefore (31) implies (30).

When $A \equiv 1 \pmod 4$, let $p_n = 2ADn + AD - 1$. For some value of n, by Dirichlet's theorem, p_n is a prime p. Since $p \equiv 1 \pmod 4$ and $p \equiv -1 \pmod A$, we have $(-A/p) = +1$. Therefore, there exists an odd integer r, such that

$$1 + Ar^2 \equiv 0 \pmod{p}, \qquad D + ADr^2 \equiv 0 \pmod{4p}. \tag{32}$$

On setting $c = 2n + 1$, $s = 2$, we note that $4p = 4Mc - s^2$, and therefore (32) implies (30).

THEOREM 7. *Let AD be without square factor and let there exist integers b, c, r, s with r odd and s even, such that*

$$D + ADr^2 + bs^2 = 4ADbc. \tag{33}$$

Then A is odd, and if D is odd, $A \equiv 3 \pmod 4$ and there exists a factor $q > 1$ dividing D, such that $(-A/q) = -1$. If D is even and $A \equiv 7 \pmod 8$, there exists a factor $q > 1$ dividing D, such that $(-A/q) = -1$.

Let $D = 2^i\delta$, δ odd, $i = 0$ or 1, $s = 2s_1$, $m_0 = ADc - s_1^2$, and suppose $q = \text{g. c. d.}(\delta, m_0)$, $\delta = q\delta_0$, $m_0 = qm$, $s_1 = qs_0$. Then $4ADc - s^2 = 4qm$, where m is relatively prime to δ and by (33),

$$2^i\delta_0(1 + Ar^2) \equiv 0 \pmod{4m}. \tag{34}$$

Write $m = 2^j\mu$, where μ is odd. Since m is prime to δ_0, evidently, for $\mu > 1$, $1 + Ar^2 \equiv 0 \pmod\mu$ and $(-A/\mu) = +1$. Since $A \equiv 3 \pmod 4$, $(A/\mu) = (-1)^{(\mu-1)/2} \cdot (\mu/A)$, and $(-1/A) = -1$. Since $m = 2^j\mu \equiv -q \pmod A$, we have

$$\begin{aligned}(\mu/A) &= (2^j/A)(m/A) \\ &= (-1)^{(q+1)/2}(-1/q)(2^j/A)(-A/q) = -(2^j/A)(-A/q).\end{aligned}$$

Hence,

$$(-A/q) = -(2^j/A)(-1)^{(\mu-1)/2}(-1/\mu)(-A/\mu) = -(2^j/A).$$

But, if $A \equiv 7 \pmod 8$, then $(2/A) = +1$; and if $i = 0$, $A \equiv 3 \pmod 8$, then, by (34), m is odd, since $\delta_0(1 + Ar^2) \equiv 4 \pmod R$. Therefore, $(2^j/A) = +1$ and $(-A/q) = -1$.

Lastly, if $\mu = 1$, then $2^j = 2^i A\delta_0 c - qs_0^2$, so that $-2^j q \equiv (qs_0)^2 \pmod A$; and, therefore, $(-2^j q/A) = +1$. But, since $A \equiv 3 \pmod 4$,

$$(-A/q) = (-1/q)(-1)^{(q-1)/2}(q/A) = -(-q/A) = -(2^j/A) = -1.$$

7. *Summary.* By Theorem 4, equations (22) are satisfied, except, possibly, when $\mu_1 = 1$ and $M_1 = AD_1$. But, by Theorem 5, the second of equations (22), with s odd, holds for every $M_1 = AD_1$ when $P > 1$. Therefore, equations (22) can be satisfied with s odd in all cases except when $M = AD$ and $(-A/p) = +1$ for every odd prime p dividing d. By Theorem 6, equations (22) can be satisfied with s even for $M = AD$ whenever A is odd and incongruent to 7 modulo 8. Theorem 7 implies that, when $A \equiv 7 \pmod 8$, if equations (22) are not satisfied with s odd, then they also fail to hold for s even. Hence, by Theorem 3, we obtain the final theorem.

THEOREM 8. *Express any positive integers a and d in their unique forms*

$$a = 4^l \lambda^2 M, \qquad d = 4^g \gamma^2 D$$

where $\lambda\gamma$ is odd, M and D are without square factor. If $l \geqq g$, the set $\Sigma(d)$ of positive non-classic ternaries represents all integers a except those for which $M = AD$, $A \equiv 7 \pmod 8$ and $(-A/p) = +1$ for every odd prime p dividing d. If $l < g$, the set $\Sigma(d)$ represents all integers a except those for which $M = AD$ and $(-A/p) = +1$ for every odd prime p dividing d.

COROLLARY. *Every set $\Sigma(d)$ represents no integer of the form $4^g D(8dm - 1)$. No set of forms $\Sigma(d)$ represents all positive integers.*

This last result offers new proof of the well known fact that no positive ternary represents all positive integers.

IOWA STATE COLLEGE,
AMES, IOWA.

IDEAL WARING THEOREM FOR THE POLYNOMIAL $m(x^3-x)/6-m(x^2-x)/2+x$.[1]

By Alvin Sugar.

This paper is numbered in sequel to a previous paper by the author [2] in which there was obtained the second known universal Waring theorem for a polynomial with a parameter.[3]

Theorem. *Every positive integer is a sum of* $m+3$ *values of* $P(x)=m(x^3-x)/6+x$ *for non-negative integers* x, *where* $m \geqq 16$.

Since the completion of that paper, Dickson [4] has proved the ideal universal Waring theorem for n-th powers, for $n>6$. This is, then, the third universal Waring theorem for a polynomial with a parameter. In this paper we shall add a fourth theorem to this list. And, furthermore, this result, as were the three preceding, is also an ideal result.

5. *A universal Waring theorem obtained by a transformation.* If we subject

$$f(x)=m(x^3-x)/6-m(x^2-x)/2+x,\ x \text{ integral and } \geqq 0, \tag{13}$$

to the transformation $x=y+1$, we get

$$f(x)=P(y)+1. \tag{14}$$

From (14) and the above theorem, we can immediately conclude that every positive integer $N>m+3$ is a sum of $m+3$ values of $f(x)$, when $m \geqq 16$. Since every integer $\leqq m+3$ is surely a sum of $m+3$ values (0 and 1 being values of $f(x)$), we have proved the first part of the following theorem.

Theorem 8. *Every positive integer is a sum of* $m+3$ *values of* (13) *for* $m \geqq 16$ *and is a sum of nine values for* $1 \leqq m \leqq 6$.[5]

---

[1] Presented to the Society, August 31, 1936. Received by the Editors, October 19, 1936.

[2] *American Journal of Mathematics*, vol. 58 (1936), pp. 783-790.

[3] The first theorem of this kind was proved by Cauchy, *Oeuvres* (2), vol. 6, pp. 320-353, and had been stated earlier by Fermat. It was only recently that James conjectured the existence of similar theorems for cubic polynomials, *American Journal of Mathematics*, vol. 56 (1934), p. 305.

[4] *American Journal of Mathematics*, vol. 58 (1936), pp. 530-535.

[5] It is known that nine values of $P(x)$ suffice for $1 \leqq m \leqq 6$. See Dickson, *Transactions of the American Mathematical Society*, vol. 36 (1934), p. 739, Theorem 12.

This theorem, however, is not an ideal Waring theorem. But, fortunately, we can obtain the ideal Waring theorem for this polynomial.

Since the constant C_1 of Theorem 1 was taken more than 10 greater than the Dickson-Baker-Webber constant, we have

THEOREM 9. *For $m \geqq 7$, every integer $\geqq 10^{12}m^{10}$ is a sum of nine or ten values of* (13) *according as the congruence $m \equiv 6 \pmod 9$ does not or does hold.*

6. *The ideal for $f(x)$.* We list a set of the first twelve values of $f(x)$:

$0, 1, 2, a = m + 3, b = 4m + 4, c = 10m + 5, d = 20m + 6, e = 35m + 7,$ $f = 56m + 8,\ g = 84m + 9,\ h = 120m + 10,\ i = 165m + 11.$

The ideal $g(f)$ is the smallest value of s for which it is true that every positive integer is a sum of s values of f. We can prove

$$B(f) = [(m+1)/2] + 3 \leqq g(f).$$

For, we see that the integer

$$\left[\frac{4m+4}{m+3}\right](m+3) - 1$$

requires at least $[(m+1)/2] + 3$ values of $f(x)$ when $m \geqq 5$. Our next step will be to prove $g(f) \leqq B(f)$ (whence $g(f) = B(f)$) by proving that every positive integer is a sum of $B(f)$ values of $f(x)$.

7. *Ascension theory.* If we replace x by $y + 1$ in (14) we get the following identity in y:

$$f(y+1) \equiv P(y) + 1. \tag{15}$$

If we write

$$d(a) = f(a+1) - f(a),$$

we see by (15) that the equality

$$d(a+1) \equiv F(a) \tag{16}$$

holds identically in a, where $F(a)$, as we recall, was defined to be

$$P(a+1) - P(a).$$

From (15) it also follows that

$$f(a+2) = P(a+1) + 1 > P(a+1). \tag{17}$$

Now from the analogue for $f(x)$ of Theorem 2, and these two relations, (16) and (17), we see that we can use the following theorems in making ascensions with the polynomial $f(x)$.

THEOREM 10. *Let every integer n, $c < n \leqq g$, be a sum of $k-1$ values of $f(x)$, and let a be an integer $\geqq 0$ for which $F(a) < g-c$. Then every integer N, $c < N \leqq g + P(a+1)$, is a sum of k values of $f(x)$.*

THEOREM 11. *Let every integer n, $c < n \leqq c + pm + q$, be a sum of k values of $f(x)$, and let t be a real number $\geqq 1$ which satisfies the inequality $F(3t+1) < pm + q$. Then every integer N,*

$$c < N \leqq (9/2)t^{2(3/2)^s} m$$

is a sum of $k + s$ values of $f(x)$.

8. *Two Lemmas.*

LEMMA 5. *Every positive integer $\leqq 20m + 6$ is a sum of $B = B(f)$ values of f for $m \geqq 17$.*

LEMMA 6. *Every integer between $20m + 6$ and $217m + 32$ is a sum of B values of f for $m \geqq 36$.*

In proving Lemma 5 we proceed by a method analogous to that employed in obtaining Lemma 2 and get the following set of numbers, so constructed that between each consecutive pair, B values suffice.

$3a = 3m + 9$, $b = 4m + 4$, $a + b = 5m + 7$, $2a + b = 6m + 10$, $6a + m - 7 = 7m + 11$, $6a + m - 6 = 7m + 12$, $3a + b = 7m + 13$, $2b = 8m + 8$, $a + 2b = 9m + 11$, $c = 10m + 5$, $a + c = 11m + 8$, $2a + c = 12m + 11$, $4a + 2b + m - 8 = 13m + 12$, $3a + c = 13m + 14$, $b + c = 14m + 9$, $a + b + c = 15m + 12$, $3a + 3b + m - 8 = 16m + 13$, $4b = 16m + 16$, $8a + 2b + m - 17 = 17m + 15$, $3a + b + c = 17m + 18$, $2b + c = 18m + 13$, $8a + c + m - 15 = 19m + 14$, $a + 2b + c = 19m + 16$, $d = 20m + 6$.

In order to prove Lemma 6 we construct the following set of intervals, over each of which $[(m-1)/2]$ values suffice. They will overlap for $m \geqq 36$. Hence for $m \geqq 36$ we can conclude that every integer n,

$$20m + 5 < n \leqq 23m + 12, \tag{18}$$

is a sum of $[(m-1)/2]$ values of f.

$(d = 20m + 6,\ 21m + 3)$, $(5b = 20m + 20,\ 21m + 9)$, $(a + d = 21m + 9,\ 22m + 4)$, $(5a + 4b = 21m + 31,\ 22m + 12)$, $(2a + d = 22m + 12,\ 23m + 5)$, $(6a + 4b = 22m + 34,\ 23m + 13)$.

We note that the interval (18) has been so selected that an ascension is unnecessary; for, in virtue of Theorems 10 and 11 (since Theorems 10 and 11 for $f(x)$ are identical with Theorems 2 and 4 for $P(x)$), we see that the ascension made in establishing Lemma 3 from (6) is valid here. And in the future, whenever an ascension is necessary we shall be careful to select our interval and the number of values which suffice over this interval in such a way that a new ascension will not be necessary i. e. in such a way that a corresponding ascension of the previous paper will suffice.

9. *The major ascensions.* The following intervals have been so constructed that every integer lying in anyone of them is a sum of $B - 11$ values of f. These intervals will overlap for $m \geqq 36$. Hence $B - 11$ values will suffice from $165m + 11$ to $171m + 14$.

$(i = 165m + 11,\ 166m - 6)$, $(c + e + h = 165m + 22,\ 166m + 1)$, $(b + 3e + f = 165m + 33,\ 166m + 8)$, $(a + 4d + g = 165m + 36,\ 166m + 9)$, $(a + 4c + 2d + g = 165m + 44,\ 166m + 13)$, $(a + b + 2c + 4e = 165m + 45,\ 166m + 14)$, $(a + i = 166m + 14,\ 167m - 5)$, $(a + c + e + h = 166m + 25,\ 167m + 2)$, $(a + b + 3e + f = 166m + 36,\ 167m + 9)$, $(2a + 4d + g = 166m + 39,\ 167m + 10)$, $(2a + b + d + 4e = 166m + 44,\ 167m + 13)$, $(a + b + 3c + 2d + e + f = 166m + 49,\ 167m + 16)$, $(a + 4c + d + 3e = 166m + 50,\ 167m + 17)$, $(2a + i = 167m + 17,\ 168m - 4)$, $(2a + c + e + h = 167m + 28,\ 168m + 3)$, $(2a + b + 3e + f = 167m + 39,\ 168m + 10)$, $(3a + 4d + g = 167m + 42,\ 168m + 11)$, $(3a + b + d + 4e = 167m + 47,\ 168m + 14)$, $(3a + 4c + 2d + g = 167m + 50,\ 168m + 15)$, $(3a + b + 2c + 4e = 167m + 51,\ 168m + 16)$, $(2a + b + 3c + 2d + e + f = 167m + 52,\ 168m + 17)$, $(2a + 4c + d + 3e = 167m + 53,\ 168m + 18)$, $(2g = 168m + 18,\ 169m - 1)$, $(3a + c + e + h = 168m + 31,\ 169m + 4)$, $(2b + 2c + f + g = 168m + 35,\ 169m + 8)$, $(a + 2b + 2c + d + e + g = 168m + 43,\ 169m + 12)$, $(a + 2b + 4c + e + g = 168m + 47,\ 169m + 14)$, $(4a + b + d + 4e = 168m + 50,\ 169m + 15)$, $(b + i = 169m + 15,\ 170m - 4)$, $(b + c + e + h = 169m + 26,\ 170m + 3)$, $(2b + 3e + f = 169m + 37,\ 170m + 10)$, $(a + 2b + d + 4e = 169m + 45,\ 170m + 14)$, $(2b + 3c + 2d + e + f = 169m + 50,\ 170m + 17)$, $(b + c + 6d + e = 169m + 52,\ 170m + 18)$, $(a + b + i$

$= 170m + 18,\ 171m - 3)$, $(a + b + c + e + h = 170m + 29,\ 171m + 4)$, $(a + 2b + 3e + f = 170m + 40,\ 171m + 11)$, $(2a + b + 2c + 3d + g = 170m + 47,\ 171m + 14)$.

Now, from the first ascension of the previous paper and from Lemmas 5 and 6, we have the following theorem.

THEOREM 12. *Every positive integer is a sum of B values of* (13) *for* $36 \leqq m \leqq 1950$.

We again follow the procedure of the previous paper, and beginning from an arbitrary point, construct a set of intervals such that

$$B - r - 3 = \left[\frac{m+1}{2}\right] - r$$

values will suffice over each, where we take

$$r = (R - A - \epsilon - 21)/2, \text{ and positive,} \tag{19}$$

and ϵ is 0 or 1 according as $R - A - 1$ is even or odd.

We begin with $f(A+1) = Rm + A + 1$, and our first interval is

$$\big(Rm + A + 1,\ (R+1)m + A - 2r\big).$$

The rest of the intervals are

$$\big((R + t - 1)m + 3R + 3t - 3,\ (R + t)m + R + t - 2r\big), \tag{20}$$

$$\big((R + t)m + R + t - 2r,\ (R + t + 1)m + R + t - 4r - 23\big), \quad (t = 1, \cdots, 10). \tag{21}$$

In constructing the intervals (20) and (21) we take for the initial point of (20) the integer $(R + t - 1)a$. The value for r in (19) was obtained by requiring that r be the greatest integer satisfying

$$T(A,t) = (R + t)m + R + t - 2r \geqq f(A+1) + ta, \quad (t = 1, \cdots, 10), \tag{22}$$

uniformly in t. Substituting the value for r from (19) in $T(A,t)$ (the end point of (20)), we get

$$T(A,t) = f(A+1) + ta + (10 - t)2 + \epsilon.$$

Hence $T(A,t)$ is a sum of 12 values of $f(x)$ (of which one value will be zero if $R - A - 1$ is even). From this we obtain the interval (21) in the usual way.

An inspection gives us the information that these intervals will overlap for

$$m \geqq Q(A) = 4R - 2A - 1 = \frac{2}{3}(A^3 - 4A) - 1.$$

Expressing r in terms of A, we have

$$r = \frac{1}{12}(A^3 - 7A) - \frac{\epsilon + 21}{2}.$$

LEMMA 7. *For $m \geqq Q(A)$, $A \geqq 6$, every integer n,*

$$Rm + A + 1 \leqq n \leqq (R + 10)m + A + 31,$$

is a sum of $[(m+1)/2] - r$ values of $f(x)$.

The analogues to statements (S_1), (S_2), and (S_3) for the polynomial now under consideration are:

(S'_1) For $Q(A) \leqq m < Q(A+1)$ every integer $> f(A+1)$ is a sum of B values, provided $A \geqq 10$.

(S'_2) For $m \geqq Q(A)$, B values will suffice from $f(A+1)$ to $f(A+2)$, when $A \geqq 10$.

(S'_3) For $m \geqq Q(A)$, every integer n, $f(11) \leqq n \leqq f(A+2)$, will be a sum of B values.

As before (S'_3) is obtained from (S'_2), and (S'_1) and (S'_3) are sufficient to prove the following theorem.

THEOREM 13. *Every integer $\geqq f(11)$ is a sum of B values of $f(x)$ for $m \geqq Q(10) = 639$.*

In proving (S'_1) we proceed as we did before. From Lemma 7 and the ascension that established (5), we have that every integer N,

$$f(A+1) \leqq N \leqq c_3 m = \frac{9}{2}(10)^{2(3/2)^r} m, \tag{23}$$

is a sum of B values of $f(x)$ for $m \geqq Q(A)$. It is evident that

$$10^{12} m^{10} \leqq c_3 m, \text{ for } Q(A) \leqq m \leqq M = (10^{-12} c_3)^{1/9}. \tag{24}$$

For $A \geqq 10$ (henceforth we shall require $A \geqq 10$) we have that

$$r > e_1 = \frac{A^3}{16}. \tag{25}$$

From (4_1) we see that

$$\frac{2}{9}\left(\frac{3}{2}\right)^{e_1} - \frac{12}{9} > e_2 = \frac{10^{-4}A^6}{2}, \tag{26}$$

and from (4_2) we see that

$$10^{e_2} > 10^{-13}A^{18} > Q(A+1). \tag{27}$$

Therefore $M > Q(A+1)$ from (23), (24), (25), (26) and (27); and we have proved (S'_1). It is evident that

$$c_3 m > Mm > mQ(A) > f(A+2),$$

and therefore by (23) we have established (S'_2).

From Lemmas 5 and 6 and Theorems 12 and 13 we get

THEOREM 14. *Every positive integer is a sum of* $[(m+1)/2]+3$ *values of* $m(x^3-x)/6 - m(x^2-x)/2 + x$ *for non-negative integers* x, *where* $m \geqq 36$.

We embody the results of Theorems 8 and 14 in a final, recapitulative theorem.

THEOREM 15. *Every positive integer is a sum of* $[(m+1)/2]+3$, $m+3$, *or* 9 *values of* $m(x^3-x)/6 - m(x^2-x)/2 + x$ *for non-negative integers* x *according as* $m \geqq 36$, $35 \geqq m \geqq 16$, *or* $6 \geqq m \geqq 1$.

THE UNIVERSITY OF CALIFORNIA,
BERKELEY, CALIFORNIA.

PROOF OF A THEOREM OF LEHMER.[1]

By PAUL M. HUMMEL.

1. Introduction. Galois first proved the theorem that if the partial quotients of a convergent, purely periodic binary continued fraction were inverted (their order reversed), the new continued fraction converged to a number belonging to the same quadratic field as the original. The like is not usually true for continued fractions of higher order as may be seen by taking almost any example at random. Lehmer [2] gave a generalization of Galois' theorem for the ternary continued fraction but remarked that his method of proof could not handle the general, or n-ary, case. The purpose of this article is to give a short elementary proof of Lehmer's theorem for the n-ary continued fraction.

2. Continued fractions of order n. The n-ary continued fraction is defined as follows: let $(x_1', x_1'', \cdots, x_1^{(n-1)})$ be any $n-1$ real numbers; by an n-ary continued fraction representation of this set we shall mean the $(n-1)$-uple sequence of numbers

$$(p_1', p_1'', \cdots, p_1^{(n-1)}; \; p_2', p_2'', \cdots, p_2^{(n-1)}; \cdots; \; p_k', p_k'', \cdots, p_k^{(n-1)}; \cdots)$$

obtained from the recursion formulas

$$(1) \qquad \begin{aligned} x_k' &= p_k' + 1/x_{k+1}^{(n-1)}, \\ x_k^{(i)} &= p_k^{(i)} + x_{k+1}^{(i-1)}/x_{k+1}^{(n-1)}, \qquad (i = 2, \cdots, n-1), \end{aligned}$$

where the $p_k^{(i)}$ are chosen by some definite law of selection.

The set of numbers $(p_k', p_k'', \cdots, p_k^{(n-1)})$ is called the k-th *partial quotient set* and is denoted for brevity by p_k. Similarly

$$x_k = (x_k', x_k'', \cdots, x_k^{(n-1)})$$

is called the k-th *complete quotient set.*

In order to apply matric methods, the complete quotients are replaced by homogeneous notation. Thus we set

[1] Received July 23, 1936.

[2] D. N. Lehmer, *Bulletin of the American Mathematical Society* (1931), pp. 565-570.

$$x_k^{(i)} = u_k^{(i+1)}/u_k', \qquad (i = 1, \cdots, n-1).$$

The defining equations (1) are then replaced by the equivalent ones

$$\begin{aligned} u_k' &= u_{k+1}^{(n)}, \\ u_k^{(i+1)} &= u_k^{(i)} + p_k^{(i)} u_{k+1}^{(n)}, \qquad (i = 1, \cdots, n-1). \end{aligned}$$

These equations may be considered as a unimodular transformation whose matrix is

$$T_k = \begin{Vmatrix} 0 & 0 & 0 \cdots & 1 \\ 1 & 0 & 0 \cdots & p_k' \\ 0 & 1 & 0 \cdots & p_k'' \\ \cdot & \cdot & \cdot \; \cdot \; \cdot & \cdot \\ 0 & 0 & 0 \cdots & p_k^{(n-1)} \end{Vmatrix}.$$

Since the matrix of a product of transformations is the product of the matrices, it follows that

$$u_1^{(r)} = \sum_{i=1}^{n} A_{k-n+i}^{(r)} u_{k+1}^{(i)}, \qquad (r = 1, \cdots, n),$$

where the matrix

$$M_k = (A_{k-n+s}^{(r)}) = T_1 T_2 \cdots T_k.$$

Since $M_{k+1} = M_k T_{k+1}$, we obtain the new set of recursion formulas

$$A_{k+1}^{(r)} = A_{k-n+1}^{(r)} + A_{k-n+2}^{(r)} p'_{k+1} + \cdots + A_k^{(r)} p_{k+1}^{(n-1)}, \qquad (r = 1, \cdots, n),$$

with initial values, $(A_{-n+s}^{(r)}) = I$.

The set of numbers $(A_k''/A_k', A_k'''/A_k', \cdots, A_k^{(n)}/A_k')$ is by definition, the k-th *convergent set*. In case the $\lim_{k\to\infty} (A_k^{(i)}/A_k')$, $(i = 2, \cdots, n)$, exists, the expansion is said to be *convergent* and it is readily shown [3] that these limits are $x_1^{(i-1)}$ respectively.

In case the partial quotient sets ultimately become periodic, the representation is said to be *periodic*. Let

$$(p_1; \cdots; \overline{p_{l+1}; \cdots; p_{l+h}}) \tag{2}$$

be a periodic n-ary continued fraction of period h with l non-recurring partial quotient sets. We define

$$P = (P_{rs}) = T_{l+1} T_{l+2} \cdots T_{l+h}.$$

[3] See O. Perron, *Mathematische Annalen*, vol. 64 (1907), pp. 2-76, or my Dissertation, The Ohio State University (1935).

to be the *period matrix* of the continued fraction. The characteristic equation of P will also be called the characteristic equation of the continued fraction. It is easily shown [4] that if the continued fraction is convergent, the $x_1^{(i)}$ belong to the algebraic field defined by the characteristic equation.

3. Linear continued fractions. If the partial quotients contained in the period of a periodic continued fraction are representable in terms of a linear parameter, the continued fraction is called *linear*.[5] Thus (2) is a linear continued fraction provided

$$p_k^{(i)} = a_i t_k + b_i, \qquad (i = 1, \cdots, n-1;\ k = l+1, \cdots, l+h).$$

An interesting property of this type of continued fraction is given by

LEHMER'S THEOREM.[6] *The characteristic equation of a linear n-ary continued fraction is unaltered by inverting the partial quotient sets contained in the period.*

To prove the theorem, we define an auxiliary matrix A as follows: the first row of A = the first column of $A = (a_1, a_2, \cdots, a_{n-1}, 0)$; the last row of A = the last column of $A = (0, a_1, a_2, \cdots, a_{n-1})$; while the remaining elements of A are determined by the formula

$$a_{r,s-1} = a_{r-1,s} + D_{s-1,r-1}, \qquad (r = 2, \cdots, n-1;\ s = 3, \cdots, n), \tag{3}$$

where $D_{ij} = (a_i b_j - a_j b_i)$. Thus each row is determined in terms of the preceding row and the D's.

We will first show that the matrix A is symmetric. Interchanging r and s in (3) gives, (since $D_{ij} = -D_{ji}$)

$$a_{s-1,r} = a_{s,r-1} + D_{s-1,r-1}. \tag{4}$$

The left members of (3) and (4) will be equal if $a_{r-1,s} = a_{s,r-1}$, that is, the r-th row will equal the r-th column provided the $(r-1)$-st row equals the $(r-1)$-st column and since the first row equals the first column, A is symmetric.

We next show that the determinant of A, $|A|$, which is a polynomial in the a_i and the b_i, is not identically zero. This is easily seen by taking all the a_i and b_i except a_1 equal to zero. This gives

---

[4] See Perron or my Dissertation, *loc. cit.*

[5] So named by Lehmer, *loc. cit.*

[6] Lehmer, *loc. cit.*, proved the theorem for $n = 3$ and 4 but stated that he had not yet found a general proof.

$$A = \begin{Vmatrix} a_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & a_1 \\ 0 & 0 & 0 & \cdots & a_1 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & a_1 & \cdots & 0 & 0 \\ 0 & a_1 & 0 & \cdots & 0 & 0 \end{Vmatrix}.$$

Whence $| A | = \pm a_1^n$, which proves the assertion.

Now consider the matric product $S_k = AT_k{}^T$. ($T^T = T$-transpose). Written at length, this is

$$S_k = \begin{Vmatrix} a_1 & a_2 & \cdots & a_{n-1} & 0 \\ a_2 & a_{22} & \cdots & a_{2,n-1} & a_1 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n-1} & a_{n-1,2} & \cdots & a_{n-1,n-1} & a_{n-2} \\ 0 & a_1 & \cdots & a_{n-2} & a_{n-1} \end{Vmatrix} \cdot \begin{Vmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & p_k' & p_k'' & \cdots & p_k{}^{(n-1)} \end{Vmatrix},$$

where, of course, $p_k{}^{(i)} = a_i t_k + b_i$. We will show that S_k is symmetric. It is apparent that the first row of S_k = the first column of $S_k = (0, a_1, a_2, \cdots, a_{n-1})$. For $(r, s = 2, \cdots, n)$ the elements e_{rs} of S_k are

$$e_{rs} = a_{r,s-1} + a_{r-1}(a_{s-1}t_k + b_{s-1}),$$

and hence

$$e_{rs} - e_{sr} = a_{r,s-1} - a_{s,r-1} + D_{r-1,s-1},$$

which in view of (3) and the fact that A is symmetric, is equal to zero. Therefore S_k is symmetric so that

$$AT_k{}^T = T_k A, \qquad (k = l+1, \cdots, l+h). \tag{5}$$

Now let the partial quotients in the period of (2) be inverted. The new period matrix is

$$P^* = T_{l+h} T_{l+h-1} \cdots T_{l+1}.$$

Multiplying on the right by the auxiliary matrix A gives

$$P^*A = T_{l+h} \cdots T_{l+2}(T_{l+1}A) = T_{l+h} \cdots T_{l+3}(T_{l+2}A)T^T{}_{l+1},$$

and by repeated use of (5), we obtain $P^*A = AP^T$. Whence

$$(P^* - \lambda I)A = A(P^T - \lambda I),$$

where λ is an indeterminate. Taking determinants gives

$$| P^* - \lambda I | \cdot | A | = | A | \cdot | P^T - \lambda I |.$$

Since $|A|$ is a polynomial in the a_i and b_i which is not identically zero, it may be divided out, giving

$$|P^* - \lambda I| = |P^T - \lambda I|,$$

and this relation holds identically when the a_i, b_i, and λ are considered as independent indeterminates. Hence the characteristic function of P^T, and therefore P, is the same as that of P^*. This proves the theorem.

4. Conclusions. There are two other obvious types of continued fractions for which the characteristic equation is invariant under inversion of the partial quotients. First, those in which the partial quotient sets read the same backward and forward. Second, those which have just two sets of partial quotients in the period. This latter statement is true because of the well known matric theorem that the characteristic equation of a product of matrices is invariant under any cyclic permutation of the matrices. It also follows indirectly from Lehmer's theorem.

Whether there are types other than the three listed for which the characteristic equation is invariant under inversion is not known.

THE UNIVERSITY OF ALABAMA.

REPRESENTATION OF LARGE INTEGERS BY CUBIC POLYNOMIALS.[1]

By MARY HABERZETLE.

By use of the prime number theory we are able to prove the following theorem regarding the representation of all sufficiently large integers. Complete proof of this theorem and the lemmas is to be found in my dissertation at the University of Chicago. Only the theorem is proved here.

Consider the ternary forms:

$$A = x_1^2 + x_2^2 + x_3^2, \quad B = x_1^2 + x_2^2 + 2x_3^2, \quad C = x_1^2 + x_2^2 + 5x_3^2,$$
$$D = x_1^2 + 2x_2^2 + 2x_3^2, \quad E = x_1^2 + 2x_2^2 + 3x_3^2, \quad F = x_1^2 + 2x_2^2 + 4x_3^2,$$
$$G = x_1^2 + 2x_2^2 + 6x_3^2, \quad H = x_1^2 + 2x_2^2 + 5x_3^2.$$

Denote a general one of these by $f = h_1x_1^2 + h_2x_2^2 + h_3x_3^2$.

THEOREM. *Let $m = 1, 2, 3, 4,$ or 5, and let t be a given positive integer prime to 6 or 30. For t prime to 6, let h_1, h_2, h_3 be given by any one of the forms, $A, B, \cdots, G$, and for t prime to 30, let h_1, h_2, h_3 be given by any one of the forms, $A, B, \cdots, H$. Then all sufficiently large integers are represented by*

$$n = ma^3 + tb^3 + \sum_{i=1}^{3} h_i(c_i^3 + d_i^3).$$

We make use of two lemmas.

LEMMA 1. *If p is a prime $\equiv 2 \pmod 3$, and if m is any integer not divisible by p, every integer not divisible by p is congruent modulo p^n to the product of a cube by m.*

A proof of Lemma 1 has been given by L. E. Dickson.[2]

LEMMA 2. *If t is a positive integer prime to 6, every integer $\geqq 23^3t$ is represented by $t\gamma^3 + 6f$, where $\gamma \geqq 0$ and $f = A, \cdots, G$. If t is a positive integer prime to 30, every integer $\geqq 23^3t$ is represented by $t\gamma^3 + 6f$, where $\gamma \geqq 0$ and $f = A, \cdots, H$.*

[1] Received August 10, 1936; revised October 29, 1936.

[2] L. E. Dickson, "Simpler proofs of Waring's theorem on cubes, with various generalizations," *Transactions of the American Mathematical Society*, vol. 30 (1928), p. 2.

Proof of Theorem. Let $w = 2(h_1 + h_2 + h_3)$, and let r be the real ninth root of $(w + m + \xi t)/(w + m + \eta t)$, where ξ is a positive number such that $m + \xi t \leqq 6$, and where $0 < \eta < \xi$. Then $r > 1$. It is known that the number of primes $\equiv 2 \pmod 3$ which exceed x and are $\leqq rx$ increases indefinitely with x. Choose as x the first radical in (1). Then for all sufficiently large integers n, there exist at least ten primes p such that

$$(1) \quad [n/(w + m + \xi t)]^{1/9} < p \leqq [n/(w + m + \eta t)]^{1/9}, \quad p \equiv 2 \pmod 3.$$

The product of the ten primes exceeds $[n/(w + m + \xi t)]^{10/9}$ and hence exceeds n if $n > (w + m + \xi t)^{10}$. Hence, not all ten are divisors of n. In what follows, let p be a prime $> m$, not dividing n, and satisfying (1). By Lemma 1, there exist integers a and M satisfying

$$n \equiv ma^3 \pmod{p^3}, \qquad n - ma^3 = p^3 M, \qquad 0 < a < p^3.$$

By (1),

$$(w + m + \eta t)p^9 \leqq n < (w + m + \xi t)p^9.$$

Therefore,

$$(w + m + \eta t)p^9 - mp^9 < n - ma^3 = p^3 M,$$
$$p^3 M < n < (w + m + \xi t)p^9.$$

Cancellation of the factor p^3 gives

$$(w + \eta t)p^6 < M < (w + m + \xi t)p^6.$$

Write $M = N + wp^6$. Then

$$\eta t p^6 < N < (m + \xi t)p^6 \leqq 6p^6.$$

For n sufficiently large, $\eta p^6 \geqq 23^3$. Then $N > 23^3 t$. By Lemma 2, N can be represented by $t\gamma^3 + 6f$, $f = h_1 x_1^2 + h_2 x_2^2 + h_3 x_3^2$, $\gamma \geqq 0$. Then

$$n = ma^3 + t(p\gamma)^3 + 6fp^3 + wp^9$$
$$= ma^3 + t(p\gamma)^3 + \sum_{i=1}^{3} h_i[(p^3 + x_i)^3 + (p^3 - x_i)^3].$$

If any $|x_i| \geqq p^3$, $N \geqq 6p^6$, contrary to the above that $N < 6p^6$. Therefore, each cube is $\geqq 0$. This proves the theorem.

THE UNIVERSITY OF CHICAGO.

GROUPS OF ORDER 64 WHOSE SQUARES GENERATE THE FOUR GROUP.[1]

By G. A. MILLER.

1. *General theorems.* When the squares of the operators of a group generate the four group its order is obviously of the form 2^m and the number of these groups whose commutator subgroup is of order 2 is known to be $7m/2 - 12$ when $m > 2$ is even and $7(m-1)/2 - 9$ when $m > 3$ is odd. The order of the commutator subgroup of such a group cannot exceed 4 since the quotient group with respect to the subgroup generated by the squares is abelian. When the order of this commutator subgroup is 4 it is the four group and it appears in the central of the group. The determination of all the groups which have the four group both for the group of their squares and also for their commutator subgroup presents many difficulties which remain unsolved. The order of such a group is at least 32 and it is known that there are 9 such groups of this order.[2] The present article is devoted to the determination of some general properties of these groups by means of which all such groups of order 64 can be readily determined. These properties seem to exhibit a better method for the determination of all the groups of order 64 than the one used in an earlier volume of the present journal.[3]

Let G represent a group which has the four group both for the group of its squares and also for its commutator subgroup. When the order of G is 32 then it contains an abelian subgroup of order 16 and each of its other operators has four conjugates. Hence none of these operators appears in an abelian subgroup of order 16. Whenever the order of G exceeds 32 then every one of its operators appears in such a subgroup. This theorem applies to all the groups which have the four group as the group generated by their squares since this four group is then in the central of the group and the corresponding quotient group is the abelian group of type 1^n. To prove the theorem in question let s represent any operator of the group and consider the subgroup composed of the operators which are commutative with s. This subgroup is at least of order 16. If it is of this order it includes the central of G and

[1] Received July 18, 1936.

[2] G. A. Miller, *Proceedings of the National Academy of Sciences*, vol. 22 (1936), pp. 112-115.

[3] G. A. Miller, vol. 52 (1930), pp. 617-636.

hence it is abelian. If it is of larger order it includes an abelian subgroup of order 16 which contains s.

When G is of order 64 all of its operators appear in five subgroups of order 16 which have the group of the squares in common and satisfy the condition that no two of them have any other common operator. This results from the fact that the abelian group of order 16 and of type 1^4 has the property that all of its operators appear in five subgroups and hence every two of these subgroups have only the identity in common. When the order of G exceeds 64 it contains a subgroup of order 64 which satisfies the given conditions. The first of these five subgroups can be chosen arbitrarily and the second can then be any of those which has only the group of the squares in common with the first. To prove that these five subgroups can be so chosen that at least two of them are abelian it is only necessary to prove that we can use for the second an abelian group after the first has been so chosen that it is abelian. When the first of these is contained in an abelian subgroup of index 2 this is obvious.

When the first of these subgroups is not contained in an abelian subgroup of index 2 we may consider the abelian subgroup of order 16 which includes an operator not contained in the first of these subgroups. If this second abelian subgroup has only four operators in common with the first no further proof is necessary. If it has eight operators in common with it the two generate a group of order 32 whose central is of order 8 and which involves exactly three abelian subgroups of order 16. The abelian subgroup of order 16 which contains an operator which is not found in this subgroup of order 32 has 8 operators in common with this subgroup of order 32. These common operators may be assumed to include the group of the squares of G. If they do not constitute the central of this subgroup of order 32 the first abelian subgroup of order 16 can clearly be so chosen as to have only four operators in common with the second.

In the other possible case G involves an abelian subgroup of index 2 since its central is of order 8, in view of the theorem which will be proved in § 3 that a necessary and sufficient condition that a group of order 64 whose commutator subgroup is the four group contains an abelian subgroup of index 2 is that its central is of order 8. This proves the following theorem: *Every group of order 2^m, $m > 5$, which has the four group for the group of its squares contains two abelian subgroups of order* 16 *which have only this four group in common.*

2. *Groups of order* 64 *which contain only five abelian subgroups of order* 16. Unless the contrary is stated it will always be assumed in what follows

that both the commutator subgroup and the group of the squares of G is the four group. When G is of order 64 it is generated by the two abelian subgroups of order 16 noted at the close of the preceding paragraph. Suppose that every operator of G which is not contained in the subgroup generated by its squares has four conjugates and hence that no two abelian subgroups of order 16 have more than four common operators. In this case G contains five and only five abelian subgroups of order 16 and these subgroups include all the operators of G and every two of them have only the group of the squares in common. It should be noted that G contains also other sets of five subgroups of order 16 such that they involve all the operators of G and that no two of them have any operators in common besides those which appear in the commutator subgroup but each of these sets includes non-abelian subgroups of order 16.

There is obviously one and only one such G in which the two generating abelian subgroups are both of type 1^4. This contains 36 operators of order 4 and each of its three other abelian subgroups of order 16 is of type 2^2. There is also one and only one such G which is generated by two abelian subgroups of types 1^4 and $2, 1^2$ respectively. It also contains 36 operators of order 4 and two of its other abelian subgroups of order 16 are of type $2, 1^2$ while the remaining one is of type 2^2. When one of the generating abelian subgroups is of type 1^4 it is not possible that each of the others is of type 2^2 since some of the remaining operators are clearly of order 2. Hence there are two and only two such groups of order 64 in which a pair of generating abelian subgroups of order 16 includes the group of type 1^4.

Suppose now that G contains no abelian subgroup of type 1^4 but that at least one of its abelian subgroups is of type $2, 1^2$. Since not all of its remaining operators could be of order 4 it results that such a G contains at least two abelian subgroups of type $2, 1^2$ and hence it is generated by two such subgroups. When the operators of order 4 in these two subgroups have a common square and operators of order 2 contained therein transform each other into themselves multiplied by this square there results a G which involves 52 operators of order 4. Its other three abelian subgroups of order 16 are therefore of type 2^2. When they transform each other into themselves multiplied by a different square there results a G which involves 44 operators of order 4. Two of its other three abelian subgroups of order 16 are of type $2, 1^2$ and their operators of order 4 have the same square but this square is not the same as the square of the operators of order 4 in the two given generating abelian subgroups of order 16.

When G contains 60 operators of order 4 each of its five abelian subgroups

of order 16 is of type 2^2. An operator of one of two such generating subgroups could not transform all the operators of the other into their inverses since G would then contain more than three operators of order 2. For the same reason an operator of one of these subgroups must transform at least one operator of order 4 in the other into its inverse. It must therefore transform four of these operators into their inverses and the other two sets of four into themselves multiplied by the common square of the other set of four. Hence the transformation is completely determined and there is one and only one such group of order 64. Hence *there are five and only five groups of order* 64 *which separately satisfy the conditions that both their commutator subgroup and the group generated by their squares is the four group and that each of their other operators has four conjugates.*

3. *Groups which contain an abelian subgroup of index* 2. When G is the direct product of a group of order 32 which has the four group both for the group of its squares and for its commutator subgroup it contains an abelian subgroup of index 2 but the 9 groups which satisfy this condition will be excluded from the considerations of the present section. Hence we shall not consider here the case when this abelian subgroup is of type 1^5. It was noted above that a necessary and sufficient condition that G contains an abelian subgroup of index 2 is that its central is of order 8. The necessity of this condition results directly from the fact that the order of the commutator subgroup of G is 4. Moreover, when the central of G is 8 and it contains a non-abelian subgroup of index 2 this subgroup has a commutator subgroup of order 2 and hence some of the operators of G which are not found in this subgroup are commutative with 16 of its operators. In all the other groups considered in the present article the central is therefore identical with the commutator subgroup.

The abelian subgroup of index 2 contained in G is characteristic since it is the only abelian subgroup of index 2 contained in G. When it is of type $2, 1^3$ then its central is of type $2, 1$ and hence there is one and only one such group. It contains 40 operators of order 4. When the abelian subgroup of index 2 is of type $2^2, 1$ the central of G is again of type $2, 1$ and two co-sets with respect to it are similar and involve only operators of order 4. The third co-set involves operators which have the same square as those of the central. There are two groups when the operators of this co-set are transformed into themselves multiplied by the square of an operator of order 4 in the central. One of these contains 40 operators of order 4 and 24 of these have a common square while the other contains 56 operators of order 4 composed of 8, 24, 24

operators respectively which have a common square. A third group results when the operators of the said co-set are transformed into themselves multiplied by an operator of order 2 which is not the square of an operator of order 4 in the central. This involves 48 operators of order 4. Hence, excluding direct products, *there are four groups of order* 64 *which involve an abelian subgroup of index* 2 *and have the four group both for their commutator subgroup and also for the group of their squares.*

4. *Some properties of the remaining groups.* In each of the remaining groups every pair of generating abelian subgroups of order 16 which have only the central of G in common must satisfy the condition that at least four operators of each of these subgroups have two and only two conjugates under the other since not every operator which is not in the central of G can have four conjugates. If one of these two subgroups has only four such operators the other will also contain only four such operators and the operators of G which have two conjugates will generate an abelian subgroup of order 16 involving 12 such operators. Each of these two generating abelian subgroups will then contain 8 operators which have four conjugates under the other and G involves exactly seven abelian subgroups of order 16. In the second possible case each of these two generating abelian subgroups contains 8 operators which have only four conjugates under the other. As four of these must give rise to a commutator which differs from the commutator to which the other four give rise each of these generating subgroup then contains 4 and only 4 operators which have eight conjugates under the other. The operators which have only two conjugates under G in this case generate G and such a G contains exactly nine abelian subgroups of order 16.

In order to simplify the considerations which follow it is desirable to note that with the exception of the groups which involve 60 operators of order 4 each of the remaining groups contains a pair of generating abelian subgroups of order 16 such that at least one of them is of type $2, 1^2$. To prove this theorem it is only necessary to prove that such a G contains at least one abelian subgroup of type $2, 1^2$ which involves the commutator subgroup. For if another abelian subgroup of order 16 would have eight operators in common with this one the two would generate a group of order 32 whose commutator subgroup would be of order 2. An abelian subgroup of order 16 containing an operator not found in this subgroup of order 32 would therefore have only the commutators of G in common with the given group of type $2, 1^2$. If G did not contain such a group its operators of order 2 which are not found in its central would appear in one or more subgroups of type 1^4. Hence it is only

necessary to prove that if G contains a subgroup of this type it must also contain one of type $2, 1^2$.

If G contains two subgroups of type 1^4 which have only the central of G in common then each of its four abelian subgroups of order 16 which has only the central in common with one of them contains four different operators from one of the three co-sets of G with respect to this subgroup, viz., from one which involves 12 operators of order 4. It therefore results that G involves an abelian subgroup of type $2, 1^2$. If it contains only one subgroup of type 1^4 then not all of the remaining operators can be of order 4 and hence there is again a subgroup of type $2, 1^2$. It therefore follows that *every group of order* 64 *which does not involve* 60 *operators of order* 4 *but has the four group for its commutator subgroup and also for the group of its squares contains the abelian subgroup of type* $2, 1^2$ *unless each of its operators which is not in its central has four distinct conjugates.*

5. *Groups in which the operators which have two conjugates generate an abelian subgroup of order* 16. The operators of two of the co-sets of G with respect to one of a pair of generating abelian subgroups have four conjugates under this subgroup while the operators of the third of these co-sets have only two conjugates under the same subgroup but half of these operators have four conjugates under G. The four abelian groups of order 16 which have only the central in common with the given group of this order have eight operators in common in pairs. Four of these belong to one of the co-sets with respect to the given abelian subgroup of order 16 but their operators belonging to the other two of these co-sets are distinct. If one of the two generating abelian subgroups of order 16 is of type $2, 1^2$ and the second is of type 1^4 the resulting G contains 32 operators of order 4 and the three additional abelian subgroups which have only the central in common with the first are of type $2, 1^2$. This G is also generated by two abelian subgroups of type 1^4 and hence there is only one such group.

When G involves an abelian subgroup of type 1^4 but no such subgroup which has only the central in common with it, it contains 48 operators of order 4 and these have only two distinct squares in sets of 24. Hence there are two and only two groups which come under the present heading and in which the operators having two conjugates generate the abelian group of type 1^4. When the operators which have two conjugates generate the abelian group of type $2, 1^2$ and G involves a second abelian subgroup of this type which has 8 operators in common with it these 8 operators may constitute the group of type 1^3 or the group of type $2, 1$. In the former case there are two groups

which involve separately 48 operators of order 4. In one of these 32 operators of order 4 have the same square while in the other at most 24 operators have this property. When the 8 operators of the abelian group of order 16 and of type $2, 1^2$ which is generated by the operators which have two conjugates and appear in the given abelian subgroup of type $2, 1^2$ constitute the group of type 2, 1 it may be assumed that at least one of the four generating abelian subgroups of order 16 which have only the central in common with the given one of type $2, 1^2$ is either of type 1^4 or of type $2, 1^2$.

When the first of these two conditions is satisfied there are two groups involving 32 and 40 operators of order 4 respectively. When this condition is not satisfied none of the remaining four generating abelian subgroup in question is of type 1^4 but at least one of them is of type $2, 1^2$. There are three such groups involving 40, 48, 48 operators of order 4 respectively. Only one of the last two involves 32 operators which have a common square. Hence there are 7 such distinct groups of order 64 which satisfy the condition that the operators which have two conjugates generate the abelian group of type $2, 1^2$, and 2 in which these operators generate the abelian group of type 1^4. It remains to determine those in which these operators generate the abelian group of type 2^2 and hence the given subgroup of type $2, 1^2$ involves four operators of order 4 which have two conjugates under the group, and all the operators which have two conjugates are of this order. There are three such groups involving 40, 48, 56 operators of order 4 respectively. Since a group in which the operators which have two conjugates generate an abelian group of order 16 could not contain 60 operators of order 4 *there are* 12 *and only* 12 *groups of order* 64 *which have the four group both for their commutator subgroup and for the group of their squares and satisfy the condition that their operators which have two conjugates generate an abelian subgroup of order* 16.

6. *Groups which are generated by their operators which have two conjugates.* Since the four group is assumed to be the central of each of these groups each of a pair of generating abelian subgroups of order 16 involves 8 operators which have two and only two conjugates under the other. There are 8 other operators in G which have two and only two conjugates while the remaining 36 operators of G have four conjugates. For each of the 9 abelian subgroups of order 16 contained in G there are 4 others which have four operators in common with it and 4 others which have eight such common operators. We shall first determine all these groups which involve at least one abelian subgroup of type 1^4.

There is one and only one group in which a pair of generating abelian subgroups of order 16 is composed of groups of type 1^4. It contains 28 operators of order 4. Two sets of 12 of these operators have a common square but the squares of one set are not the same as those of the other. The remaining 4 have the third commutator of order 2 for their common square. When G contains at least one abelian subgroup of type 1^4 there is a pair of generating abelian subgroups of order 16 which includes it. We proceed to determine those groups which involve one such subgroup but not a pair of such generating subgroups. Two of the co-sets of the group with respect to this subgroup of type 1^4 are conjugate. If we multiply the operators of these co-sets successively by operators of order 4 whose squares appear therein there result two conjugate groups of order 64 involving 36 operators of order 4 but if we multiply the operators of these co-sets by an operator of order 4 having a different square there results a group of order 64 which has 44 operators of this order. Hence *there are three such groups of order* 64 *which involve separately* 9 *abelian subgroups of order* 16 *and include such a subgroup of type* 1^4.

It remains to determine the groups which do not contain an abelian subgroup of type 1^4. Near the close of § 2 it was noted that there is one and only one group of order 64 which contains 60 operators of order 4 and in which every operator of this order has four conjugates. Each of the other groups determined above involves less than 60 operators of order 4. Any additional groups which belong to the category of groups considered in the present article and involve 60 operators of order 4 must therefore come under the present section. We proceed to determine these groups.

A pair of abelian generating subgroups of order 16 in such a group which have only the central in common must satisfy the conditions that each of them is of type 2^2 and that each involves two independent generators such that one of them is commutative with one of the other having a different square and transforms the second of the other into its inverse. As the co-set which involves the former of these generators involves operators of order 4 having two distinct squares the group of order 32 thus obtained is completely determined. A second independent generator of the former of these subgroups can be chosen in two essentially different ways and hence there are two such groups which involve 60 operators of order 4. In one of these 36 operators of order 4 have the same square while in the other at most 28 of these operators satisfy this condition. Hence *there are three and only three groups of order* 64 *which satisfy the conditions that they separately involve* 60 *operators of order* 4 *and have the four group both for the group of their squares and for their commutator subgroup.* Only two of these three groups belong to the present section.

Since the remaining groups which belong to the present section cannot involve an abelian subgroup of type 1^4 and cannot involve 60 operators of order 4 they must involve an abelian subgroup of type $2, 1^2$ and the four other abelian subgroups which have only the central in common with this subgroup are of one or more of the following two types: $2, 1^2$; 2^2. Suppose first that G is generated by two such subgroups of type $2, 1^2$ whose operators of order 4 have the same square. Not all of their operators of order 2 can be commutative since these groups are supposed to contain no abelian subgroup of type 1^4. There are two such groups in which all the operators of one of these subgroups are transformed into their inverses by an operator of the other. One of these contains 44 operators of order 4 while the other contains 52 such operators. Each of these groups contains four other abelian subgroups of type $2, 1^2$ and the operators of order 4 in these have the same square as in the two given generating subgroups of this type.

Suppose that G is generated by two abelian subgroups of type $2, 1^2$ whose operators of order 4 have the same squares and that one of these subgroups again contains an operator which is commutative with the operators of order 2 in the other but which does not transform into their inverses the operators of order 4 in this other. There are three such groups. Two of these involve 44 operators of order 4 while the third involves 52 such operators. The last of these groups can easily be distinguished from the one noted in the preceding paragraph which contains the same number of operators of order 4 by the fact that the latter contains 36 such operators having a common square while the former contains at most 28 such operators. To distinguish the three given groups involving 44 operators of order 4 it may be noted that in one of these there are at most 20 operators of order 4 which have a common square while in each of the others there are 28 such operators. The latter can easily be distinguished by the fact that in one of them all the operators of order 4 in the abelian subgroups of type $2, 1^2$ have a common square while in the other these operators have two different squares.

We have now determined the five groups which are generated by two abelian subgroups of type $2, 1^2$ such that one of these subgroups contains operators which are commutative with the operators of order 2 in the other but not with its operators of order 4 and that the operators of order 4 in these generating subgroups have a common square. If the two generating abelian subgroups of order 16 satisfy these conditions with the exception that operators of order 4 of the one are commutative with operators of order 4 of the other there result two additional groups. One of these can readily be distinguished from those which precede by the fact that it involves 36 opera-

tors of order 4. The second involves 44 such operators but only four abelian subgroups of type $2, 1^2$ while the preceding group which contains 44 operators of order 4 involve a larger number of such subgroups. Hence *there are seven and only seven groups of order* 64 *which separately satisfy the three conditions that both the group of their squares and their commutator subgroup is the four group, that each of them is generated by its operators which have two conjugates and also by two abelian subgroups of type* $2, 1^2$ *whose operators of order* 4 *have a common square.*

If a group is generated by two abelian subgroups of type $2, 1^2$ whose squares are different but not by two such subgroups whose squares are the same it may be assumed that these two squares are fixed. There is one such group in which all the operators of order 2 in one of these generating subgroups are invariant under 8 operators of the other. This group contains 52 operators of order 4 and four abelian subgroups of type $2, 1^2$ but those whose operators of order 4 have the same square have more than 4 common operators. There is also one such group in which there are operators of order 4 in one of these subgroups which are commutative with 8 operators of the other. This group also contains 52 operators of order 4 and four abelian subgroups of type $2, 1^2$. These appear in two sets of two each such that these operators of order 4 have a common square but each such pair has 8 operators in common.

It remains to consider the groups which separately involve no two generating abelian subgroups of type $2, 1^2$ but involve at least one subgroup of this type. There is one such group. It contains 52 operators of order 4 and two abelian subgroups of type $2, 1^2$ but these have eight operators in common and hence they do not generate G. Each of its remaining seven abelian subgroups of order 16 contained therein is of type 2^2.

UNIVERSITY OF ILLINOIS.

NECESSARY AND SUFFICIENT CONDITIONS FOR GENERATING CERTAIN SIMPLE GROUPS BY TWO OPERATORS OF PERIODS TWO AND THREE.[1]

By ABRAHAM SINKOV.

1. *Introduction.* Necessary and sufficient conditions for generating the simple groups of order 60, 168, and 660 by two operators of periods two and three are already known.[2] Professsor G. A. Miller [3] has shown that no such definition is possible for the simple group of order 360. The purpose of the present study is to determine necessary and sufficient conditions for the two remaining simple groups G_{504} and G_{1092}, whose orders do not exceed 1092.

In addition to the above-mentioned groups, G_{5616} has been studied by K. E. Bisshopp [4] who showed that there are only two abstractly distinct ways in which this group may be generated by two operators of periods two and three. However, each of his abstract definitions required nine distinct restrictions upon the generators to insure sufficiency and apparently no attempt was made to study their independence. It will be shown in the present study that five conditions are sufficient to define the group completely in either case.

2. *General relations.* The various definitions which will be obtained in this study will all be based on the relations $S^3 = T^2 = (ST)^n = (S^{-1}T^{-1}ST)^p = 1$, which we shall designate $(2, 3, n; p)$. However, instead of employing the generators S and T, we shall introduce the following substitution: [5]

$$P = (ST)^{-1} \qquad S = P^2Q$$
$$Q = (ST)^2S \qquad T = P^3Q.$$

[1] Received April 22, 1936.

[2] H. R. Brahana, "Pairs of generators for the known simple groups whose orders are less than one million," *Annals of Mathematics*, vol. 31 (1930), pp. 543-544. This paper will hereafter be referred to as "Pairs of generators."

[3] G. A. Miller, *Bulletin of the American Mathematical Society*, vol. 7 (1900-1901), p. 426.

[4] *Bulletin of the American Mathematical Society*, vol. 37 (1931), p. 99.

[5] This same substitution for $(ST)^7 = 1$ was used by Professor Brahana in his paper "Certain perfect groups generated by two operators of periods two and three," *American Journal of Mathematics*, vol. 50 (1928), p. 349. This paper will hereafter be referred to as "Certain perfect groups."

Since $Q = S^{-1}(S^{-1}TST)S$, its period is the same as that of the commutator of S and T. Hence the relations $(2, 3, n; p)$ will be replaced by

$$P^n = Q^p = (QP^3)^2 = (QP^2)^3 = 1.$$

The set of relations $(2, m, n; p)$, i. e.

$$S^m = T^2 = (ST)^n = (S^{-1}T^{-1}ST)^p = 1$$

is intimately connected with the relations

$$A^m = B^n = C^p = (AB)^2 = (BC)^2 = (CA)^2 = (ABC)^2 = 1$$

which have been studied by Dr. H. S. M. Coxeter.[6] He has designated the group defined by these relations $G^{m,n,p}$ and has shown that if any one of the numbers m, n, p is even (say p) then $G^{m,n,p}$ contains as a subgroup of index two the group defined by $(2, m, n; p/2)$. Should two of the numbers be odd (say n and p) then, by putting $Q = C^2$; $P = A$, he is enabled to express the relations $G^{m,n,p}$ in the form

$$P^n = Q^p = (Q^{(p+1)/2}P)^2 = (QP^2)^m$$
$$= \{(QP^2)^{(m-1)/2}P\}^2 = \{Q^{(p+1)/2}(QP^2)^{(m-1)/2}\}^2 = 1.$$

These relations are not all independent; in fact any one of the last three may be considered redundant. Putting $m = 3$, we may write

$$G^{3,n,p} \equiv (2, 3, n; p); \qquad (Q^{(p+1)/2}P)^2 = 1$$

and observe that either one of the relations $(QP^3)^2 = 1$, $(QP^2)^3 = 1$ may be replaced by $(Q^{(p+3)/2}P^2)^2 = 1$.

For convenience in some of the manipulation which will be required throughout this paper, we will now establish some consequences of the relations $(2, 3, n; p)$.

$$QP^2Q = P^{-2}Q^{-1}P^{-2} = P \cdot P^{-3}Q^{-1} \cdot P^{-2}$$

hence

$$QP^2Q = PQP \tag{1}$$
$$P^2QPQ^a = P \cdot PQP \cdot Q^a = PQP^2Q^{a+1}$$
$$= Q \cdot P^2QPQ^a \cdot Q = Q^{-a} \cdot P^2QPQ^a \cdot Q^{-a}$$
$$= Q^{-a}P^{-1} \cdot P^3Q \cdot P = Q^{-a}P^{-1}Q^{-1}P^{-2}$$

whence

$$(P^2QPQ^a)^2 = 1 \quad \text{and also} \quad (PQP^2Q^{a+1})^2 = 1. \tag{2}$$

[6] These results, which have been communicated to me by Dr. Coxeter, have not yet been submitted for publication.

From (1)

$$QP^{-1} = P^{-2}Q^{-1}PQ = P \cdot P^{-3}Q^{-1} \cdot PQ$$
$$= PQP^4Q \tag{3}$$
$$P^3Q \cdot P^{-1} = P^4QP^4Q = Q^{-1}P^{-4}$$
$$(QP^4)^3 = 1. \tag{4}$$

From (2)

$$P^2Q^2 \; PQP \cdot PQ^2PQ = 1$$
$$P^2Q^2P^2Q \cdot PQP \cdot Q = 1$$
$$(Q^2P^2)^3 = 1. \tag{5}$$

These results lead to some new properties in connection with the independence of the relations defining $G^{3,n,p}$ (n, p odd). We have from (2)

$$QPQ^{(p-1)/2}P^2 = P^{-2}Q^{-[(p-1)/2]}P^{-1}Q^{-1} = P^{-2}Q \cdot Q^{-[(p+1)/2]}P^{-1} \cdot Q^{-1}.$$

If we assume $(Q^{(p+1)/2}P)^2 = 1$, the right member becomes $P^{-2}QPQ^{(p-1)/2}$. Hence $P^2 \cdot QPQ^{(p-1)/2} \cdot P^2 = QPQ^{(p-1)/2}$. If now, the period of P is an odd number, this leads to

$$QPQ^{(p-1)/2} = PQPQ^{(p-1)/2}P = PQP \cdot Q^{(p-1)/2} \cdot PQ^{(p+1)/2} \cdot Q^{-[(p+1)/2]}$$
$$= QP^2Q \cdot Q^{(p-1)/2}Q^{-[(p+1)/2]}P^{-1}Q^{-[(p+1)/2]} = QPQ^{-[(p+1)/2]}$$

whence $Q^p = 1$.

Therefore, the relations $(2, 3, n)$; $(Q^{(p+1)/2}P)^2 = 1$ imply $Q^p = 1$ and give a complete definition of $G^{3,n,p}$ provided n and p are both odd. These relations are independent whenever $n > 5$, $p > 5$.

In the same way, starting with the operator $QPQ^{(p+3)/2}P^2$, it can be shown that the relations $(2, 3, n; p)$; $(Q^{(p+3)/2}P^2)^2 = 1$ imply $(Q^{(p+1)/2}P)^2 = 1$.

3. *The simple group of order* 504. We proceed first to determine the number of different ways in which G_{504} may be generated by two operators of periods two and three.[7] The 63 operators of period two are all conjugate [8] and it is therefore sufficient to consider only one of them, say

$$T \equiv (1, 3)(2, 6)(4, 5)(7, 8)$$

as a possible generator. The largest subgroup of G within which T is invariant is abelian, of order eight and type $(1, 1, 1)$. It is generated by T and the two operators

---

[7] The general procedure to be followed has been outlined by Professor Brahana in "Pairs of generators," pp. 542, 543.

[8] The properties of G_{504} and G_{1092} which are made use of here as well as the method of representing these groups as permutation groups are given in Dickson's *Linear Groups with an Exposition of the Galois Field Theory* (1901), chap. XII.

$$T' \equiv (1,4)(2,8)(3,5)(6,7); \qquad T'' \equiv (1,7)(2,5)(3,8)(4,6).$$

Under this subgroup, the 56 operators of period three are divided up into seven sets of conjugate operators; the members of each set satisfy with T the same abstract relations. We give below one operator from each set and note the period of its product with T.

	Period of ST
$S_1 = (2,5,3)(4,7,6)(1,9,8)$	9
$S_2 = (3,6,4)(5,1,7)(2,9,8)$	9
$S_3 = (4,7,5)(6,2,1)(3,9,8)$	7
$S_4 = (5,1,6)(7,3,2)(4,9,8)$	9
$S_5 = (6,2,7)(1,4,3)(5,9,8)$	7
$S_6 = (7,3,1)(2,5,4)(6,9,8)$	7
$S_7 = (1,4,2)(3,6,5)(7,9,8)$	2

The permutation $D \equiv (1,2,4)(3,6,5)$ which corresponds to the transformation $Z' = Z^2$ transforms G_{504} into itself according to an outer automorphism. It is commutative with T and transforms S_1 into S_2, S_2 into S_4. Hence the relations satisfied by T with S_1 are the same as those satisfied by T with either S_2 or S_4. Similarly D transforms S_3 into S_6, S_6 into S_5. Hence these three operators satisfy the same relations with T. Since T and S_7 generate a dihedral group of order 6, it follows that there are only two posssible definitions for G_{504}.

Now, it has already been shown by Professor Brahana [9] that the relations $(2,3,7;9)$; $(Q^6P^2)^2 = 1$ are sufficient to define G_{504}. It follows from the last result in Art. 2 that these relations imply $(Q^5P)^2 = 1$. Hence $G_{504} \equiv G^{3,7,9}$. This result enables us to set down at once a complete definition for both cases.

THEOREM. *Two operators of periods two and three generate G_{504} if and only if they satisfy one of the following sets of independent relations*

$$A: \quad (2,3,7); \ (Q^5P)^2 = 1$$
$$B: \quad (2,3,9); \ (Q^4P)^2 = 1.$$

It can be shown that the condition $(Q^5P)^2 = 1$ in A may be replaced by $(Q^6P^2)^2 = 1$. Hence the relations used by Professor Brahana to define G_{504} are not independent.

4. *The simple group of order* 1092. The 91 operators of period two con-

[9] "Certain perfect groups," p. 354.

tained in G_{1092} are all conjugate and it is again sufficient to consider only one of them. We choose it to be

$$T \equiv (1, 12)(2, 11)(3, 10)(4, 9)(5, 8)(6, 7).$$

The largest subgroup of G within which T is invariant is the dihedral group of order 12 generated by

$$(1, 4, 3, 12, 9, 10)(2, 8, 6, 11, 5, 7)$$

and

$$(1, 12)(2, 6)(3, 4)(7, 11)(9, 10)(13, 14).$$

Under this subgroup the 180 remaining operators of period three are divided up into 16 conjugate sets. Of these, two contain only six distinct operators each; the remaining 14 sets each contain 12 operators. As in the case of G_{504}, we give below one member of each of these sets together with the period of its product with T.

	Period of ST
$S_1 = (1, 13, 10)(2, 3, 6)(4, 9, 11)(5, 12, 7)$	6
$S_2 = S_1^2$	6
$S_3 = (2, 10, 4)(11, 13, 5)(3, 6, 7)(8, 12, 9)$	6
$S_4 = S_3^2$	6
$S_5 = (3, 10, 12)(4, 6, 13)(5, 11, 8)(14, 9, 7)$	3
$S_6 = S_5^2$	3
$S_7 = (2, 11, 14)(3, 4, 8)(5, 9, 10)(6, 13, 7)$	2
$S_8 = (1, 13, 12)(9, 4, 14)(3, 8, 6)(5, 10, 7)$	2
$S_9 = (2, 8, 9)(4, 14, 13)(5, 10, 6)(7, 12, 11)$	7
$S_{10} = (2, 7, 8)(3, 10, 11)(5, 13, 9)(6, 12, 14)$	13
$S_{11} = (2, 3, 4)(6, 9, 11)(7, 12, 14)(8, 10, 13)$	7
$S_{12} = S_{11}^2$	7
$S_{13} = (2, 14, 5)(3, 9, 13)(4, 7, 11)(8, 10, 12)$	7
$S_{14} = S_{13}^2$	7
$S_{15} = (1, 10, 6)(3, 8, 9)(4, 11, 12)(7, 13, 14)$	7
$S_{16} = (1, 10, 4)(3, 6, 14)(5, 12, 8)(9, 13, 11)$	13

Since no group satisfying the relations $(2, 3, 6)$ is simple,[10] it is obvious that none of the first eight of the above operators when coupled with T will generate the entire group. S_{12} and S_{14} satisfy with T the same defining relations as do S_{11} and S_{13}, respectively.

S_9 is transformed into S_{15} by the substitution

[10] G. A. Miller, *Quarterly Journal*, vol. 33 (1901-1902), p. 76.

$$(1,2)(3,5)(4,7)(6,9)(8,10)(11,12)(13,14)$$

which is commutative with T. Similarly S_{10} is transformed into S_{16} by

$$(1,7,10,5,9,11,12,6,3,8,4,2)$$

which is also commutative with T. Hence there remain only S_9, S_{10}, S_{11} and S_{13} to be considered. The periods of the commutators of S_9, S_{11} and S_{13} with T are respectively 13, 6, 7. Hence there are four distinct definitions for G_{1092}.

It has already been proved that the relations $(2,3,7;6)$ and $(2,3,7;7)$ are sufficient [11] to determine G_{1092}. We pass then to the definition in terms of $(2,3,7;13)$. That these four relations are not sufficient to define G_{1092} follows from the fact that Professor Brahana has shown that the simple group of order 9828 can also be generated by two operators satisfying the same relations.[12] It becomes necessary then to consider a fifth condition, and it will appear in what follows that the new relation $(Q^2P^6)^3=1$ is sufficient to define G_{1092}. Indeed, it will be shown to be sufficiently strong to make the condition $Q^{13}=1$ redundant.

We first consider some consequences of the initial relations $(2,3,7)$; $(Q^2P^6)^3=1$. In order to facilitate the verification of the manipulation which follows, a number has occasionally been written to the right to indicate the particular relation in Art. 2 which is being used in the process of simplification.

$$\begin{aligned} Q^3P^6Q^3P^2Q^{-1}P^2 &= Q^3P^6Q^3P^2\cdot Q^{-1}P^4\cdot P^5 \\ &= Q^3P^6Q^2P^6\cdot PQP^5QP^5. \end{aligned}$$

But
$$\begin{aligned} PQP^5QP^5 &= PQP\cdot P^4QP^4\cdot P && (1),\ (4) \\ &= QP^5Q^{-1}P = QP\cdot P^4Q^{-1}\cdot P \\ &= Q\cdot PQP^4Q\cdot Q^{-1} = Q^2P^6Q^{-1} && (3) \end{aligned}$$

$$\begin{aligned} \therefore\quad Q^3P^6Q^3P^2Q^{-1}P^2 &= Q(Q^2P^6)^3Q^{-1}=1 \\ Q^3P^6Q^3 &= P^5QP^5. \end{aligned}$$

Now $P^5QP^5=P^5(QP^3)P^2$ is of period two. Hence Q^6P^6 is also of period two. Therefore

$$\begin{aligned} Q^{-6} &= P^{-1}Q^6P^{-1} = P^{-1}Q\cdot Q^4\cdot QP^{-1} && (1) \\ &= QPQ^{-1}P^{-2}\cdot Q^4\cdot P^{-2}Q^{-1}PQ \\ Q^{-8} &= PQ^{-1}P^{-1}\cdot P^{-1}Q^4P^{-2}Q^{-1}\cdot P = P^2\cdot P^{-1}Q^{-1}P^{-1}\cdot QP^2Q^{-4}P^2 && (2),\ (1) \\ &= P^2\cdot Q^{-1}P^{-2}Q^{-1}\cdot QP^2Q^{-4}P^2 = P^2Q^{-5}P^2 \\ 1 &= QP^2Q^8P^2Q^{-6} = QP^2Q\cdot Q^7P^2Q^{-6} && (1) \\ &= P\cdot QPQ^7P^2\cdot Q^{-6} = P^{-1}Q^{-7}P^{-1}Q^{-7} && (2) \\ (Q^7P)^2 &= 1. \end{aligned}$$

---

[11] "Certain perfect groups," p. 354; A. Sinkov, *Bulletin of the American Mathematical Society*, vol. 41 (1935), p. 239.

[12] "Pairs of generators," p. 532.

It now follows from the results at the end of Art. 2 that $Q^{13} = 1$ is a consequence of the initial relations. It then results from Dr. Coxeter's work that $(Q^8P^2)^2 = 1$ and that the group in question is either $G^{3,7,13}$ or a quotient group of it.

With the aid of these further relations, it becomes a simple matter to prove that the conditions $(2, 3, 7)$; $(Q^2P^6)^3 = 1$ define G_{1092}. The procedure and the notation are the same as those used by Professor Brahana [13] in his study of $(2, 3, 7; 6)$.

Every combination of P and Q is reducible to an operator in some one of the 156 co-sets obtainable from those listed below.

$$\begin{array}{lll} a = 1 & e = Q^2P & i = Q^2P^6 \\ b = QP & f = Q^2P^2 & j = Q^3P \\ c = QP^4 & g = Q^2P^4 & k = Q^3P^2 \\ d = QP^5 & h = Q^2P^5 & l = Q^3P^4. \end{array}$$

The representation of P and Q as permutations of these 156 symbols is as follows:

$$\begin{aligned} Q &= (aa_1a_2 \cdots a_{12})(bb_1b_2 \cdots b_{12}) \cdots (ll_1l_2 \cdots l_{12}) \\ P &= (abb_{12}a_{12}cdc_1)(a_2efc_{12}ghi)(a_3jkg_{12}lh_4d_{10}) \\ &(a_4j_9e_2l_{12}k_9d_8i_9)(a_5e_5k_4k_8e_7a_8c_7)(a_6b_5f_7e_6f_8b_7a_7) \\ &(a_9i_6d_5k_3l_5e_{10}j_7)(a_{10}d_3h_5l_4g_2k_{12}j_3)(a_{11}i_2h_9g_1c_2f_2e_{12}) \\ &(b_1f_1b_{11}d_{12}i_{12}i_3d_1)(b_2k_{11}e_{11}i_{11}g_{10}d_4h_{10})(b_3e_9j_2g_9l_{10}i_7h_6) \\ &(b_4k_7j_6l_9f_9c_8d_6)(b_8d_7c_6f_6l_8j_{10}k_5)(b_9h_3i_8l_7g_5j_1e_3) \\ &(b_{10}h_{12}d_9g_4i_4e_1k_1)(c_3j_4l_1i_1l_3j_{12}c_{11})(c_4h_1f_3l_2f_{12}h_8c_{10}) \\ &(c_5j_5f_4f_{11}j_{11}c_9g_7)(d_2k_2k_{10}d_{11}g_3j_8g_{11})(e_4l_{11}h_{11}l_6e_8i_{10}i_5) \\ &(f_5f_{10}h_7g_6k_6g_8h_2)(a)(b_6). \end{aligned}$$

As a corollary to this definition of G_{1092} we will now determine the order of $G^{3,7,13} \equiv (2, 3, 7)$; $(Q^7P)^2 = 1$. We shall prove first that these conditions imply $(Q^2P^6)^6 = 1$.

$$\begin{aligned} PQ^2PQ^{-2} &= PQ \cdot QP^3Q \cdot Q^{-1}P^5Q^{-2} \\ &= PQP^4Q \cdot Q^{-2}P^5Q^{-2} = QPQ^2P^2 && (3),\ (5) \\ &= Q^2P^6 \cdot PQ^{-1}P \cdot Q^2P^2 = Q^2P^6Q^{-1} \cdot P^3Q \cdot P^2 && (3) \\ &= Q^2P^6Q^{-2}P^6 \end{aligned}$$

whence $(PQ^2PQ^{-2})^2 = 1$. From this result, we have

$$\begin{aligned} P^6Q^2P^6 &= Q^2PQ^{11}PQ^2 \\ Q^2P^6Q^2P^6Q^2 &= Q^4PQ^{11}PQ^{11} \cdot Q^6 = Q^4PQ^{11}PQ^{11}PQ^7P \\ (Q^2P^6)^3 &= Q^4(Q^2P^6)^{-3}Q^9 \end{aligned}$$

[13] "Certain perfect groups," pp. 352, 353.

that is, $(Q^2P^6)^3$ is transformed into its inverse by Q^4. It is therefore commutative with Q and, as a result $(Q^2P^6)^3 = (Q^2P^6)^{-3}$; $(Q^2P^6)^6 = 1$.

We see then that G_{1092} is a factor group of $G^{3,7,13}$, obtained by adjoining the additional restriction [14] $(Q^2P^6)^3 = 1$. The invariant subgroup which gives rise to G_{1092} as a quotient group is the group generated by the complete set of conjugates involving $(Q^2P^6)^3$.

Now $(Q^2P^6)^3$ is invariant under transformation by Q and is therefore equal to $(P^6Q^2)^3$. It is also invariant under transformation by P. For, $P(P^6Q^2)^3P^{-1} = (Q^2P^6)^3$. Therefore $(Q^2P^6)^3$ is invariant in $G^{3,7,13}$, and the order of this perfect group is consequently 2184.

A complete definition for G_{1092} in terms of $(2, 3, 13; 7)$ can be readily obtained from the definition based on $(2, 3, 7; 13)$. We have already seen that G_{1092} may be defined by

$$C: \quad P^7 = Q^{13} = (QP^2)^3 = (QP^3)^2 = (Q^7P)^2 = (Q^2P^6)^3 = 1.$$

Let us now introduce the substitution $P^2 = \bar{Q}$; $Q = \bar{P}^2$ and let us require that $\bar{P}^{13} = 1$. Then $P = \bar{Q}^4$; $\bar{P} = Q^7$. The defining relations become

$$D: \quad \bar{P}^{13} = \bar{Q}^7 = (\bar{Q}\bar{P}^2)^3 = (\bar{Q}^5\bar{P}^2)^2 = (\bar{Q}^4\bar{P})^2 = (\bar{Q}^3\bar{P}^4)^3 = 1,$$

in which the roles played by P and Q in C have been interchanged. Since the additional condition $(Q^2P^6)^3 = 1$ has been replaced by $(Q^3P^4)^3 = 1$ it follows that G_{1092} is completely defined by $(2, 3, 13)$; $(Q^3P^4)^3 = 1$. We have then the following necessary and sufficient condition:

THEOREM. *Two operators of periods* 2 *and* 3 *generate the simple group of order* 1092 *if and only if they satisfy one of the following sets of independent relations*

$$A: \quad (2, 3, 7; 6)$$
$$B: \quad (2, 3, 7; 7)$$
$$C: \quad (2, 3, 7); \quad (Q^2P^6)^3 = 1$$
$$D: \quad (2, 3, 13); \quad (Q^3P^4)^3 = 1.$$

5. *The simple group of order* 5616. It is easily verified by direct calculation with the generating permutations given by Bisshopp that the period of the commutator of S and T is four for one pair of generators (the pair for which $\bar{S} = R^{-1}SR$) and six for the other $(\bar{S} = R^{-9}SR^9)$. Hence G_{5616} is a factor group of $(2, 3, 13; 4)$ and $(2, 3, 13; 6)$. We consider first the case $Q^4 = 1$ and proceed to show that one further restriction on P and Q is sufficient

[14] Dr. Coxeter has found that this additional restriction is equivalent to $(CBA)^7 = 1$, where the operators A, B, and C are the generators of $G^{3,7,13}$ in his definition.

to define G_{5616}. This additional relation is $QP^7QP^7 = P^6Q^2P^7Q^2$ which implies $(QP^9)^4 = 1$. For,

$$QP^7QP^7 = QP^7 \cdot P^{10}Q^3P^{10} \cdot P^7 = QP^4Q^3P^4$$
$$Q^3P^4Q^3P^4 = Q^2P^6Q^2P^7Q^2 = (P^7Q^2)^{-1}Q^2(P^7Q^2).$$

This in turn implies $(QP^{10})^4 = 1$. For,

$$QP^9 = QP^3 \cdot P^6 = P^{10}Q^3P^6 = P^7(P^3Q^3)P^6.$$

The period of $Q^{-1}P^7$ is also four since $Q^{-1}P^7 = Q^{-1}P^{-3} \cdot P^{10} = P^3QP^{10}$. The proof of sufficiency will consist in showing that the five relations chosen can be used to establish all of the conditions given by Bisshopp. The reduction of these conditions to expressions involving P and Q will be omitted as that involves no particular difficulty. The following reductions will however be found useful in simplifying the process.

$$\bar{S} = P^2Q^2$$
$$\bar{S}T = P^2QP^{10}$$
$$(\bar{S}T)^4 = P^3Q^2P^9QP^{12}Q$$
$$\bar{S}^2T\bar{S}T = QP^{11}Q^3P^9$$

(1) $(P^2QP^{10})^8 = 1$

$$\begin{aligned}(P^2QP^{10})^2 &= P^2QP^2 \cdot P^{10}QP^{10}\\ &= Q^3P^{11} \cdot Q^3P^{10} \cdot QP^{10}\\ &= Q^3PQ^2 \cdot P^{10}Q^3 \cdot Q\\ &= Q^3PQ^3P^3Q\\ &= (P^3Q)^{-1}P^4Q^3(P^3Q)\end{aligned}$$

(2) $QP^{11}Q^2P^4Q^3P^2Q^3P^{11}Q^2 = P^3Q^3P^2QP^4Q^3P^6Q^2P^2$

$QP^{11}Q^2P^4Q^3P^2Q^3 \cdot P^{11}Q^2P^{11}Q^2 \cdot P^7QP^9Q^3P^{11}QP^{10} = 1$ (5)

$QP^{11}Q^2P^4Q^3P^2 \cdot QP^9QP^9 \cdot Q^3P^{11}QP^{10} = 1$ (5)

$QP^{11}Q^2P^4 \cdot Q^3P^6Q^3P^6 \cdot Q^2P^5Q = 1$

$Q \cdot P^{11}Q^2P^{11}Q^2P^{11} \cdot Q = Q^4 = 1$

(3) $(QP^{11}Q^3P^9)^6 = 1$

$$\begin{aligned}(QP^{11}Q^3P^9)^3 &= QP \cdot QP^9 \cdot Q^3P^{11}Q^2P^5Q\\ &= QP^5Q^3P^4Q^3P^4 \cdot Q^2P^{11}Q^2 \cdot P^5Q \qquad (5)\\ &= QP^5Q^3P^4Q^3 \cdot P^6Q^2P^7Q^2 \cdot Q^3\\ &= QP^5Q^3P^{12} \cdot QP^3 \cdot PQP^8Q^3\\ &= (PQP^8Q^3)^{-1}QP^3(PQP^8Q^3)\end{aligned}$$

(4′)
$$P^5Q^2P^2QP^9 = P^4QP^2Q^2P^{10}$$
$$P^5Q^2P^2 \cdot QP^{12} \cdot Q^2P^{11}Q^3P^9 = 1 \qquad (3)$$
$$P^5Q^2 \cdot P^3Q \cdot P^4 \cdot Q^3P^{11}Q^3 \cdot P^9 = 1$$
$$P^5QP^3QP^{11} = P^{26} = 1$$

(5′)
$$P^5Q^2P^5Q^2P^{10}QP^5Q = PQ^3PQ^2P^{10}Q$$
$$QP^4Q^2P^5Q^2P^{10}QP^8Q^2P^{12} = 1$$
$$QP^2 \cdot P^2Q^2P^2 \cdot P^3Q \cdot QP^{10}QP^8Q^2P^{12} = 1 \qquad (5)$$
$$P^{11}Q^3P^{11}QP^{11} \cdot QP^{10}QP^{10}QP^{10} \cdot P^{11}Q^2P^{12} = 1$$
$$P^{11}Q^3P^{11}QP \cdot Q^3P^{11} \cdot Q^2P^{12} = 1$$
$$P^{11}Q^3P^{11} \cdot QP^3Q \cdot P^2Q^3P^{12} = 1$$
$$P^{11}Q^3P^{10}Q^3 \cdot P^{12} = 1$$
$$P^{26} = 1$$

(6′)[15]
$$P^5Q^3P^9Q^2P^{10}Q^2P^{11} = Q^3P^7QP^9Q$$
$$P^5Q^3P^9Q^2P^{10}Q^2 \cdot P^{11}Q^3 \cdot P^4Q^3P^6Q = 1$$
$$P^5Q^3P^9Q^2 \cdot P^{10}Q^3 \cdot P^2QP^6Q^3P^6Q = 1$$
$$P^9 \cdot P^9Q^3P^9Q^3 \cdot P^5QP^6Q^3P^6Q = 1$$
$$P^9QP^9Q \cdot P^6Q^3P^6Q = 1$$
$$Q^3P^4 \cdot Q^3P^{10}Q^3 \cdot P^6Q = 1$$
$$Q^4 = 1.$$

It follows then that G_{5616} is completely defined by the relations

$$(2, 3, 13; 4);\ QP^7QP^7 = P^6Q^2P^7Q^2.$$

In the same manner as above, it can be demonstrated that one additional restriction $Q^2P^6Q^2 = P^2Q^5P^2$ adjoined to $(2, 3, 13; 6)$ is sufficient to define G_{5616}. The manipulation is considerably more involved but the procedure is essentially the same.

We have then the following theorem:

Two operators of periods two and three generate G_{5616} if and only if they satisfy one of the following sets of relations

$$A:\ (2, 3, 13; 4);\qquad QP^7QP^7 = P^6Q^2P^7Q^2$$
$$B:\ (2, 3, 13; 6);\qquad Q^2P^6Q^2 = P^2Q^5P^2.$$

The author wishes to express here his extreme indebtedness to Dr. Coxeter for the many valuable suggestions and criticisms received in the course of this research.

WASHINGTON, D. C.

[15] N. B. There is a typographical error in the statement of this relation. The right member should be $\bar{S}(\bar{S}T)^4\bar{S}(\bar{S}T)^3\bar{S}^2$.

HYPERGROUPS.[1]

By H. S. WALL.

A hypergroup is a system in which any two elements a, b can be combined to form the product ab, which is a complex of n not necessarily distinct elements of the system. Here n is a fixed integer ≥ 1. If $n = 1$ the hypergroup reduces to an ordinary group. If $[ab]$, called the bracket product, is the set of all the distinct elements of ab, the totality of elements a such that $[ax]$ and $[xa]$ are single elements for every x forms a group with respect to the bracket product. This group is called the nucleus. A hypergroup in which the nucleus is not vacuous is related to the Mischgruppe of Loewy and Baer (cf. § 3), and is called a hypergroup of type M.

This paper is in three parts. Part I contains a postulational basis for abstract hypergroups. In Part II there is a discussion of isomorphism and automorphism for hypergroups, and the notion of conjugate sets in groups is extended to hypergroups of type M. Part III contains a study of special hypergroups, principally multiplicative systems derived from groups.

PART I.

ABSTRACT HYPERGROUPS.

1. *Postulational basis and definitions.* A *hypergroup* is a system H of distinct *elements* in which there is defined an operation called *multiplication.* With each hypergroup there is associated an integer $n \geq 1$ called the *dimension* of the hypergroup.

I. *The product postulate.* If a, b are two elements of H, not necessarily distinct, then the *product* ab is a complex of n uniquely determined elements of H. We shall write for convenience

$$ab = c_1 + c_2 + \cdots + c_n = \sum_{i=1}^{n} c_i.$$

The order in which the elements are written in the "sum" is immaterial, but ab is in general different from ba. The n elements of the product ab are not necessarily distinct.

[1] Received December 28, 1935; revised July 24, 1936.

Let $A = a_1 + a_2 + \cdots + a_p$, $B = b_1 + b_2 + \cdots + b_q$ be two complexes of elements of H. Then we shall define the product AB to be the complex

$$\sum_{i=1}^{p} \sum_{j=1}^{q} a_i b_j.$$

II. *The associative law.* If a, b, c are three elements of H, then

$$a(bc) = (ab)c = abc.$$

The members of the last equality have a meaning by the preceding definition of the product of two complexes.

III. *The identity postulate.* There is at least one element e in H, called an *identity* element, such that for every element a in H the products ae and ea both contain the element a at least once.

IV. *The postulate of the inverse.* There is an identity element e in H such that corresponding to each element a in H there is at least one element a^{-1} in H such that the products aa^{-1} and $a^{-1}a$ both contain e at least once. The element a^{-1} is called an *inverse* of a.

Evidently a hypergroup of dimension unity is a group.

The *bracket product* $[ab \cdots c]$ of two or more elements of H will be defined to be the complex of all the distinct elements in the product $ab \cdots c$. An element a of H such that $[ax]$ and $[xa]$ are single elements for every x in H will be called a *scalar*. The set of all the scalars of H will be called the *nucleus*. A hypergroup in which the nucleus contains at least one element is a *hypergroup of type M*. It will appear later that a hypergroup of type M is closely related to the *Mischgruppe* of Loewy and Baer.

We shall employ the notation nG ("n-group") for a hypergroup of dimension n; and nGM for an nG of type M. An nG_1 is an nG containing just one identity element, and in which each element a has just one inverse a^{-1} such that if aa' or $a'a$ contains e then $a' = a^{-1}$. An nG_1M is an nG_1 of type M. The notation $a \in K$, where a is an element and K is a set of elements, signifies that a is contained in K.

Since $b \in a(a^{-1}b)$ it follows that there is an element $x \in a^{-1}b$ such that $b \in ax$. Likewise, there is an element $y \in ba^{-1}$ such that $b \in ya$. We therefore have this result.

THEOREM 1. *If a, b are any two elements of an nG, H, then H contains two elements x and y such that*

$$b \in ax \quad \text{and} \quad b \in ya. \tag{1}$$

Suppose that $x_1, x_2, \cdots, x_{n+1}$ are $n+1$ different values of x such that $b \in ax$. Then

$$e \in b^{-1}ax_i, \qquad (i = 1, 2, \cdots, n+1).$$

Thus there is at least one of the n elements in the product $b^{-1}a$ which has more than one inverse.

THEOREM 2. *If in Theorem* 1 H *is an* nG_1, *then there are at most* n *elements* x *and at most* n *elements* y *satisfying* (1).

The upper bound given by this theorem for the number of solutions of (1) is actually attainable as will be seen from examples to be given later.

THEOREM 3. *Let* H *be an* nG *of finite order in which the relation* $b \in ax$ *has just one solution* x *for each pair of elements* a, b. *Then all the elements of* H *are scalars.*

Proof. If some product ac contains two or more distinct elements then, under the hypothesis, there must be an element d for which the product ad contains no elements. But this is contrary to postulate I.

From this theorem and Theorem 4 below it follows that an nG of finite order in which the relation $b \in ax$ has a unique solution x for every a and b is an ordinary group with respect to the bracket product.

THEOREM 4. *The nucleus of an* nGM *is a group with respect to the bracket product.*

Proof. Let H contain a scalar element δ. Then

$$[\delta\delta^{-1}] = [\delta^{-1}\delta] = e$$

where e is an identity element. Let x be any element of H, and determine y in accordance with Theorem 1 so that $x \in \delta y$, or, since δ is a scalar, $x = [\delta y]$ and $[\delta^{-1}x] = [ey]$. Now $[\delta ey] = [\delta y] = a$ (say) and hence $[\delta\delta^{-1}x] = [ex] = a$, so that $[ex]$ is a single element of H for every x in H. In like manner we find that $[xe]$ is a single element of H for every x in H. Suppose now that H contains another identity e'. Then $[ee'] = e = e'$. Thus *an* nGM *contains a unique identity element* e, *and* e *is a scalar.*

It follows from the preceding discussion that if δ is a scalar, then δ^{-1} is a scalar; and it is easy to see that δ can have but one inverse.

Let σ, δ be two scalars, and put $[\sigma\delta] = c$. Then

$$[\sigma^{-1}\sigma\delta] = [\sigma^{-1}c], \qquad [e\delta] = [\sigma^{-1}c], \qquad [\sigma^{-1}c] = \delta.$$

Choose y in accordance with Theorem 1 so that $x \in y\sigma^{-1}$, where x is any assigned element of H. Then

$$[y\delta] = [y\sigma^{-1}c] = [xc]$$

is a single element of H. In like manner we find that $[cx]$ is a single element of H. Thus c is a scalar. This completes the proof of the theorem.

An nGM is not necessarily an nG_1. For consider the following example.

	e	a	b
e	$e+e$	$a+a$	$b+b$
a	$a+a$	$e+b$	$e+b$
b	$b+b$	$e+a$	$e+a$

This is a $2GM$ but not a $2G_1$. It is noteworthy that $aa = ab$, $a \neq b$.

2. *Properties of hypergroups of type M.* Let H be an nGM and suppose that

$$ab_1 = ab_2 = \cdots = ab_{n+1}, \qquad b_i \neq b_j,\ i \neq j. \tag{2}$$

Then $a^{-1}ab_1 = a^{-1}ab_2 = \cdots = a^{-1}ab_{n+1}$. Now $a^{-1}ab_{n+1}$ contains n^2 elements of which n are the element b_{n+1} repeated n times. Among the remaining $n(n-1)$ elements there must be the elements $b_1, b_2, \cdots, b_n$ each repeated n times. But this is impossible. Hence we have this result.

THEOREM 5. *If H is an nGM then the relation* (2) *is impossible.*

Of course a relation obtained by reversing the orders of the factors in (2) is impossible in an nGM.

THEOREM 6. *The only element a of an nGM such that $[aa] = a$ is the identity element e.*

Proof. Put $a^{-1}a = e + c_2 + c_3 + \cdots + c_n$. Then since

$$aa = a + a + \cdots + a = na, \qquad a^{-1}aa = na^{-1}a,$$

we have

$$na + (c_2 + c_3 + \cdots + c_n)a = ne + n(c_2 + c_3 + \cdots + c_n).$$

Now if $a \neq e$, then $a = c_2$ (say) and hence

$$na + (c_3 + c_4 + \cdots + c_n)a = ne + n(c_3 + c_4 + \cdots + c_n).$$

By repeated application of this argument we shall arrive finally at the equality

$$na + c_n a = ne + nc_n,$$

which is impossible if $a \neq e$.

THEOREM 7. *If in an nGM* $[ab]$ *and* $[bc]$ *are single scalars, then* b *is a scalar.*

Proof. Determine y and z so that $x \in ya$ and $x \in cz$. Then $[xb]$ and $[bx]$ are single elements. Since x is arbitrary, b is a scalar.

COROLLARY. *If in an nGM* $[b'b] = [bb''] = e$, *the identity, then* b *is a scalar and* $b' = b'' = b^{-1}$.

THEOREM 8. *Let* a *be an element of an* nG_1M *such that* $[aa^{-1}] = e$. *Then* $[a^{-1}a] = e$.

Proof. Let x be any element of an nG_1M, H. Then by Theorem 1 there is an element y in H such that $x \in a^{-1}y$. Now $[aa^{-1}] = e$, $[aa^{-1}y] = y$, and therefore

$$[ax] = y. \tag{3}$$

Likewise we find that

$$[xa^{-1}] = z, \tag{4}$$

a single element of H. As special instances of (3) we have

$$[aa] = b, \qquad [ab^{-1}] = c, \qquad [ac] = d, \tag{5}$$

all single elements of H. Hence $[aab^{-1}] = [bb^{-1}] = [ac] = d$. Thus

$$[bb^{-1}] = e, \tag{6}$$

so that b is an element which satisfies the hypothesis of our theorem.

Put

$$\begin{aligned} b^{-1}a &= a^{-1} + t_2 + t_3 + \cdots + t_n, \\ b^{-1}b &= e + s_2 + s_3 + \cdots + s_n, \\ a^{-1}a &= e + \sigma_2 + \sigma_3 + \cdots + \sigma_n. \end{aligned}$$

Then from the first equality (5) we obtain $b^{-1}aa = nb^{-1}b$, or

$$\begin{aligned} e + \sigma_2 + \sigma_3 + \cdots + \sigma_n + (t_2 + t_3 + \cdots + t_n)a \\ = ne + n(s_2 + s_3 + \cdots + s_n). \end{aligned} \tag{7}$$

This holds if $n > 1$. Of course our theorem is granted if $n = 1$.

Assume now that the theorem is false. Then there must be an integer k, $0 < k < n$, such that for every element w satisfying the equation $ww^{-1} = ne$,

$w^{-1}w$ contains the element e at least k times; and such that there is at least one value a of w for which $a^{-1}a$ contains e just k times. Then if the notation is properly chosen, (7) takes the form

$$ke + \sigma_{k+1} + \sigma_{k+2} + \cdots + \sigma_n + (t_2 + t_3 + \cdots + t_n)a = nke + n(s_{k+1} + s_{k+2} + \cdots + s_n),$$

where $\sigma_i \neq e$, $i = k+1, k+2, \cdots, n$. But this implies that $t_i = a^{-1}$, $i = 2, 3, \cdots, n$, or, $[b^{-1}a] = a^{-1}$. Determine x so that $a^{-1} \varepsilon xb^{-1}$, apply (4), and we get

$$[xb^{-1}a] = [xa^{-1}] = z,$$

and consequently $a^{-1}a = ne$, contrary to the assumption concerning a.

COROLLARY. *If a is an element of an nG_1M such that $[aa^{-1}] = e$, then a is a scalar.*

This is a consequence of Theorems 7 and 8.

THEOREM 9. *If a, b are elements of an nG_1M such that $[ab]$ is a single scalar, then a and b are scalars.*

Proof. Determine x so that $a^{-1} \varepsilon bx$; then $[ab] = c$, $[abx] = [cx] = d$, $[aa^{-1}] = d = e$. Hence by the above corollary, a is a scalar. In like manner we find that b is a scalar.

An important consequence of this discussion concerns subhypergroups. A subset K of elements of H is a *subhypergroup* of H if the elements of K satisfy the postulates for a hypergroup. We have this theorem.

THEOREM 10. *If the identity e of an nGM H is in K, a subhypergroup of H, then K is an nGM whose nucleus is a subgroup of the nucleus of H.*

Proof. Since K contains the scalar e, K is an nGM and e is the identity of K. Let a be any scalar of K and a^{-1} an inverse of a in K. Then $[aa^{-1}] = [a^{-1}a] = e$, so that by Theorem 7 a is a scalar of H.

3. *The Mischgruppe of Loewy*[2] *and Baer.*[3] A *Mischgruppe* M is a system in which the product km is defined when and only when k is in a subset K of M called the *Kern*. The Mischgruppe may be defined by the following four postulates.

I. The Kern K is an ordinary group.

---

[2] Loewy, *Crelle*, Bd. 157 (1927), pp. 239-254.

[3] Baer, *Sitzungsbericht Heidelberger Akad. Wiss. Math-nat. Kl.* (1928) (4).

II. If a and b are any two elements of M then the product ab is defined if and only if a is in K; and when defined, ab is a uniquely determined element of M.

III. If a and b are in K, and c is in M, then

$$(ab)c = a(bc).$$

IV. If a is in K and b is in M then $ab = b$ if and only if a is the identity element of the group K.

Baer omitted the explicit statement of IV although it is clearly independent of I, II, III as the following example shows.

	e	a	A	B	C	D	E
e	e	a	A	B	C	D	E
a	a	e	D	E	C	A	B

Given a Michgruppe M which is not of order 2 nor a group. Then there is determined (not in general uniquely) a group G with subgroup E such that M is simply isomorphic with the set of all distinct cosets

$$M': \quad Eg_1, Eg_2, Eg_3, Eg_4, \cdots; \; Eh_1, Eh_2, Eh_3, \cdots$$

where $g_1, g_2, g_3, \cdots$ are commutative with E, and $h_1, h_2, h_3, \cdots$ are not commutative with E. In M' multiplication is defined by the formula

$$(Eg)(Ea) = (Ega).$$

This holds if and only if g is commutative with E: $Eg = gE$. If we put

$$(Ea)(Eb) = \sum_{x \text{ in } E} (Eaxb)$$

then with this definition of multiplication M' is a hypergroup of type M whose nucleus (with respect to the bracket product) is the Kern of the Mischgruppe. Conversely, in an nGM let us identify the nucleus with the kern K; and define the symbol " o " as follows. Let $a \circ b = [ab]$ if and only if a is in the nucleus. Then the system so defined satisfies postulates I, II, III of the Mischgruppe. Postulate IV is not in general satisfied.

We have seen that multiplication can be defined between arbitrary elements of a Mischgruppe in such a way that the new system so obtained in a hypergroup of type M. This can be done in more than one way. For consider the Mischgruppe given by the following table:

	e	a	A	B
e	e	a	A	B
a	a	e	B	A.

To this there correspond the following two hypergroups of dimension 2.

	e	a	A	B
e	$e+e$	$a+a$	$A+A$	$B+B$
a	$a+a$	$e+e$	$B+B$	$A+A$
A	$A+A$	$B+B$	$e+B$	$a+A$
B	$B+B$	$A+A$	$a+A$	$e+B$

	e	a	A	B
e	$e+e$	$a+a$	$A+A$	$B+B$
a	$a+a$	$e+e$	$B+B$	$A+A$
A	$A+B$	$A+B$	$e+a$	$e+a$
B	$B+A$	$B+A$	$a+e$	$e+a$.

The first is a $2G_1M$ while the second is not of this character.

4. *The identity groups.* Let E_a denote the set of elements of the nucleus of an nGM, H, such that $[E_a a] = a$. It is clear that E_a contains at least one element, namely the identity element; and that E_a is a group with respect to the bracket product. We shall call E_a the left a-identity group. Similarly, E'_a is the right a-identity group. It is the subgroup of the nucleus of H such that $[aE'_a] = a$. E_a and E'_a may be of different orders.

THEOREM 11. *If H is an nGM, and a any element of H, then E_a is contained in aa^{-1} and E'_a is contained in $a^{-1}a$.*

Proof. Let δ be an element of E_a. Then $[\delta a] = a$, $[\delta aa^{-1}] = [aa^{-1}]$. Hence since $\delta \in \delta aa^{-1}$ it follows that $\delta \in aa^{-1}$. Likewise we find that if δ is in E'_a then $\delta \in a^{-1}a$.

THEOREM 12. *If H is an nG_1M, and a any element of H, and we put*

$$[aa^{-1}] = E_a + F_a,$$
$$[a^{-1}a] = E'_a + F'_a,$$

then F_a and F'_a contain no scalars.

Proof. It will suffice to prove that if δ is any scalar $\in aa^{-1}$ then δ is in E_a. Write

$$aa^{-1} = e + \delta + S.$$

Then $\delta^{-1}aa^{-1} = n\delta^{-1} + ne + \delta^{-1}S$, and consequently $[\delta^{-1}a] = a$ so that δ^{-1} is in E_a. Therefore, since E_a is a group, $\delta \in E_a$, as was to be proved.

THEOREM 13. *In an nG_1M let an element δ of E_a be repeated just r times in aa^{-1}. Then every element of E_a is repeated just r times in aa^{-1}.*

Proof. By hypothesis $aa^{-1} = r\delta + S$, where S does not contain δ. Then

$$aa^{-1}\delta^{-1} = nre + T$$

where T does not contain e. Hence $a^{-1}\delta^{-1} = na^{-1}$, $naa^{-1} = nre + T$, so that aa^{-1} contains e just r times. Now if aa^{-1} contains an element δ' of E_a just s times, $s \neq r$, then the preceding argument would show that aa^{-1} contains e just s times, which is impossible.

COROLLARY. *If in Theorem* 13 E_a *contains an element repeated just* r *times in* aa^{-1}, *then for every* x *there is a* y *such that* ay *contains* x *repeated at least* r *times.*

Proof. By Theorem 13, aa^{-1} contains e repeated just r times. Hence $aa^{-1}x$ contains x at least nr times. But this is impossible unless $a^{-1}x$ contains an element y such that ay contains x at least r times.

THEOREM 14. *If* a *and* b *are elements of an* nG_1M *such that*

$$[ab] = \delta_1 + \delta_2 + \cdots + \delta_k + S \qquad (k \geq 1)$$

where $\delta_1, \delta_2, \cdots, \delta_k$ *are scalars and* S *contains no scalars, then*

$$E_a = [\delta_i\delta_1^{-1}] + [\delta_i\delta_2^{-1}] + \cdots + [\delta_i\delta_k^{-1}], \qquad (i = 1, 2, \cdots, k).$$

Proof. It follows from the hypothesis that

$$\delta_1^{-1}a = \delta_2^{-1}a = \cdots = \delta_k^{-1}a \ (= nb^{-1}).$$

Consequently $[\delta_i\delta_j^{-1}]$ is in E_a. Conversely, let δ be in E_a. Then $[\delta a] = a$, $[\delta ab] = [ab] = \delta_1 + \delta_2 + \cdots + \delta_k + S$, $[\delta_i^{-1}\delta a] = b^{-1}$. Put $[\delta_i^{-1}\delta] = \sigma$. Then $[\sigma ab] = e + T$ (say), $[ab] = \sigma^{-1} + U$, and therefore there is a j such that $\sigma^{-1} = \delta_j$, $[\delta^{-1}\delta_i] = \delta_j$, $\delta = [\delta_i\delta_j^{-1}]$, which completes the proof of the theorem.

COROLLARY. *If in an* nG_1M, E_a *contains just* k *elements, then if* $[ab]$ *contains one scalar it must contain just* k *scalars.*

Let B_a denote the set of elements of the nucleus of an nGM which are commutative with a. Clearly B_a is a group with respect to the bracket product; and any element which is in two of the groups B_a, E_a, E'_a is necessarily in the third also.

If b is in B_a, then

$$[(b^{-1}E_a b)a] = [b^{-1}E_a ab] = [b^{-1}ab] = a,$$
$$[a(b^{-1}E'_a b)] = [b^{-1}aE'_a b] = [b^{-1}ab] = a.$$

That is, $[b^{-1}E_a b] = E_a$, $[b^{-1}E'_a b] = E'_a$. In particular we have this result.

THEOREM 15. *In an nGM the elements common to E_a and E'_a form an invariant subgroup of B_a.*

Let $\Delta = E_a + [\delta_2 E_a] + [\delta_3 E_a] + \cdots$ be a decomposition of the nucleus Δ of an *nGM*, H, into distinct cosets. Then

$$[\Delta a] = a + [\delta_2 a] + [\delta_3 a] + \cdots$$

is a decomposition of the complex $[\Delta a]$ into distinct elements. If H is an nG_1M and we put $e = \delta_1$, $a_i = [\delta_i a]$, then it is easy to show that

$$[a_i^{-1}a_i] = [a_j^{-1}a_j], \qquad (i, j = 1, 2, 3, \cdots);$$
$$[\delta_i^{-1}a_i a_i^{-1}\delta_i] = [aa^{-1}], \qquad (i = 1, 2, 3, \cdots).$$

5. *Matrix representation of hypergroups.* Let $H(a_1, a_2, \cdots, a_k)$ be an nG of order k. Then there exists a set of k by k matrices in one to one correspondence with the elements of H in such a way that when

$$a_i a_j = a_{s_1} + a_{s_2} + \cdots + a_{s_n},$$

then

$$A_i A_j = A_{s_1} + A_{s_2} + \cdots A_{s_n},$$

where $A_\nu \sim a_\nu$, $\nu = 1, 2, 3, \cdots, k$.

The matrices A_ν may be formed as follows. Let $x_1, x_2, \cdots, x_k$ be positive integers, and put $X = \sum_{i=1}^{k} x_i a_i$, $X_j = a_j X = \sum_{i=1}^{k} x_i{}^j a_i$, where the $x_i{}^j$ are positive integers of the form

$$x_i{}^j = \sum_{r=1}^{k} p^j{}_{ri} x_r, \qquad (i, j = 1, 2, 3, \cdots, k),$$

in which $p^j{}_{ri} \geq 0$ is an integer. The matrices $A_j = (p^j{}_{ri})$, $j = 1, 2, 3, \cdots, k$, will be seen to have the required property. For, on the one hand,

$$a_i a_j X = (a_{s_1} + a_{s_2} + \cdots + a_{s_n})X = \sum_{r=1}^{k} X_{s_r} = \sum_{r=1}^{k} (x_r{}^{s_1} + x_r{}^{s_2} + \cdots + s_r{}^{s_n})a_r$$
$$= \sum_{r=1}^{k} x'_r a_r,$$

where $x'_r = \sum_{t=1}^{k} (p^{s_1}_{tr} + p^{s_2}_{tr} + \cdots + p^{s_n}_{tr})x_t$; while on the other hand,

$$a_i a_j X = a_i X_j = \sum_{r=1}^{k} x_r'' a_r \quad \text{where} \quad x_r'' = x_r' = \sum_{t=1}^{k} \Big(\sum_{u=1}^{k} p^i{}_{tu} p^j{}_{ur} \Big) x_t.$$

Thus we have the matrix equation $A_iA_j = A_{s_1} + A_{s_2} + \cdots + A_{s_n}$, as was to be proved.

The matrix representation is related to the following more general kind of representation suggested to me by Dr. Baer. Let H be a *finite* or *infinite* nG. Denote by H^* the set of all finite subsets of H. If a is an element of H, S an element of H^*, then denote by $f_a(S)$ the set of all the elements x of H which are contained in sets sa with s in S. Since n is finite $f_a(S)$ is finite, and $f_a(S)$ is therefore a single-valued function defined for all elements of H^* and mapping H^* into itself. If a and b are two elements of H, then

$$f_b(f_a(S)) = \sum_{x \text{ in } ab} f_x(S).$$

This representation of the hypergroup is analogous to the well-known Cayley representation of groups as permutation groups.

PART II.

Conjugate Sets and Isomorphism.

6. *Complete sets of conjugates in an nGM.* Let H be an *nGM* with nucleus Δ. Let $a_1 \in H$, and let

$$[a_1] \tag{8}$$

be the different elements of H of the form $[\delta^{-1}a_1\delta]$ where δ is in Δ. Any element of (8) is called a *conjugate* of any other element of (8). The set is the *complete set of conjugates* of any one of its members. If b_1 is any element of H not in (8) then the complete set of conjugates of b_1 has no elements in common with (8).

Theorem 16. *In an nGM let the nucleus Δ be of finite order q, and B_a, the subgroup of all elements of Δ commutative with a, of order p. Then the number of elements in the complete set of conjugates of a is q/p.*

Proof. Let $\delta_1, \delta_2, \cdots, \delta_p$ be the elements of B_a, and let δ be any element of Δ not in B_a. Then

$$[\delta_1\delta], [\delta_2\delta], \cdots, [\delta_p\delta]$$

transform a into one and the same element $a_2 \neq a$. Moreover, these are the only elements of Δ which transform a into a_2. Similarly, there are just p elements in Δ which transform a into a_3, an element $\neq a, a_2; \cdots$; and there are just p elements which transform a into a_m, an element $\neq a, a_2, a_3, \cdots, a_{m-1}$. Thus $q = mp$, $m = q/p$, as was to be proved.

7. *Conjugate subhypergroups.* If H is an nGM and K a subhypergroup of H, then by Theorem 10 if the identity e of H is in K, K is an nGM whose nucleus Σ is a subgroup of the nucleus Δ of H.

Let K be any subhypergroup of an nGM H. Let δ be an element of the nucleus Δ of H. Then

$$[\delta^{-1}K\delta] = K_2$$

is a sub-nG of H conjugate to K. The set of all the conjugates of K is a complete set of conjugate subhypergroups. Let Δ_1 be the subgroup of all elements of Δ which are commutative with K; and suppose Δ is of finite order q and Δ_1 of order p. Then by an argument similar to that used in proving Theorem 16 we find that *the number of subhypergroups in the complete set of conjugates to which K belongs is q/p.*

In case the identity of H is in K we have this theorem.

THEOREM 17. *The elements of the nucleus Δ of an nGM H which are commutative with a sub-nG K, containing the identity of H, form a subgroup Δ_1 of Δ (with respect to the bracket product) which contains the nucleus Σ of K as an invariant subgroup.*

Proof. By Theorem 10, Σ is a subgroup of Δ. Let σ be any element of Σ; and let

$$K = \Sigma + [\Sigma a_2] + [\Sigma a_3] + \cdots$$

be a decomposition of K into mutually exclusive sets. Then

$$[\sigma^{-1}K\sigma] = \Sigma + [\Sigma a_2\sigma] + [\Sigma a_3\sigma] + \cdots = K.$$

Hence Σ is a subgroup of Δ_1. Moreover, if $\delta_1 \in \Delta_1$, then since

$$[\delta_1^{-1}K\delta_1] = K,$$

it must follow that $[\delta_1^{-1}\Sigma\delta_1] = \Sigma$, so that Σ is an invariant subgroup of Δ_1.

THEOREM 18. *If δ is a scalar and $[\delta^{-1}K\delta] = K_2$ is a sub-nG conjugate to the sub-nG K of H, then the group of elements of the nucleus Δ of H which are commutative with K_2 is $[\delta^{-1}\Delta_1\delta]$ where Δ_1 is the group of elements of Δ commutative with K.*

Proof. If $[\delta^{-1}\delta_1\delta] \in [\delta^{-1}\Delta_1\delta]$, then

$$\left[[\delta^{-1}\delta_1\delta]^{-1}K_2[\delta^{-1}\delta_1\delta]\right] = [\delta^{-1}\delta_1^{-1}\delta\delta^{-1}K\delta\delta^{-1}\delta_1\delta] = [\delta^{-1}K\delta] = K_2,$$

so that this element is commutative with K_2. Conversely, if $[d^{-1}K_2d] = K_2$, $d \in \Delta$, then

$$[d^{-1}\delta^{-1}K\delta d] = [\delta^{-1}K\delta], \qquad \left[[\delta d\delta^{-1}]^{-1}K[\delta d\delta^{-1}]\right] = K,$$

and therefore $[\delta d\delta^{-1}]$ is an element δ_1 of Δ_1. Thus $d = [\delta^{-1}\delta_1\delta]$ ε $[\delta^{-1}\Delta_1\delta]$, as was to be proved.

THEOREM 19. *No sub-nGM* of an *nGM with nucleus of finite order whose nucleus is a proper subgroup of the nucleus of the main nG can contain elements belonging to every one of the complete sets of conjugates.*

Proof. Let H be the *nGM*, and suppose that such a sub-*nG*, K, can exist. Then the complete set of conjugate subhypergroups to which K belongs would contain all the elements of H. Let Δ_1 be the group of elements of Δ, the nucleus of H, commutative with K; and let $\mathfrak{Z}$ be the nucleus of K. Then the order v of Δ_1 is $\geq$ the order w of $\mathfrak{Z}$ by Theorem 17. Let N be the order of Δ. Then there are in all N/v subhypergroups in the conjugate set to which K belongs, and each of these contains w elements in its nucleus. Hence all these nuclii together cannot have more than

$$1 + w(N/v) - (N/v) = 1 + \frac{N(w-1)}{v} = 1 + \frac{Nw - N}{v} < 1 + \frac{Nw - w}{v} \leqq N$$

elements, which is impossible inasmuch as these nuclii together must make up Δ.

THEOREM 20. *The intersection of two or more sub-nG_1's of an nG_1 H, having a common identity element, is a sub-nG_1 of H.*

The proof of Theorem 20 will be omitted.

THEOREM 21. *If H is an nG_1M with finite or infinite nucleus, and K is a sub-nG_1 of H containing the identity of H, then the intersection of the complete set of conjugates to which K belongs is a normal sub-nG_1 of H i. e. is commutative with every element of the nucleus of H.*

Proof. Since K is an nG_1 containing the identity of H, it is easy to see that every one of the conjugates of K is an nG_1, and these have a common identity, namely the identity of H. Consequently, by Theorem 20, the intersection of these conjugates is a sub-nG_1 of H.

Denote by S the complete set of conjugates of K. Then since the nucleus Δ of H is a group, $[\delta^{-1}S\delta] = S$ for every $\delta \varepsilon \Delta$. Let T be the intersection of the subhypergroups in S. Then clearly $[\delta^{-1}T\delta] = T$, as was to be proved.

The following theorems are easy to verify.

THEOREM 22. *Let S be a complete set of conjugates in an nGM H. If*

S generates an nGM K, then K is a normal subhypergroup of H, and every other normal subhypergroup which contains an element of S must contain K as a subhypergroup.

THEOREM 23. *Let S be a complete set of conjugate subhypergroups in an nGM H. Then if these generate a subhypergroup K, the latter is a normal subhypergroup of H, and every other normal subhypergroup which contains one of the hypergroups of S must contain K.*

8. *Isomorphism of two hypergroups.* Two hypergroups H and H' of dimension n are *simply isomorphic* if the elements of H and H' can be put in one to one correspondence in such a way that if a and b are two elements of H which correspond to elements a' and b', respectively, of H', then the elements of the complexes ab and $a'b'$ are in one to one correspondence.

If H and H' are of type M with nuclii Δ and Δ', respectively, then Δ and Δ' are simply isomorphic groups. For, in the first place, the identity elements are in correspondence. In fact, if b' in Δ' corresponds to the identity element e in Δ, then

$$[ee] = e, \qquad [b'b'] = b'.$$

Consequently, by Theorem 6, $b' = e'$, the identity element of H'. If a' in H' corresponds to δ in Δ, then let a'^{-1} in H' correspond to δ^{-1}. Then

$$[\delta\delta^{-1}] = e \text{ implies that } [a'a'^{-1}] = e',$$
$$[\delta^{-1}\delta] = e \text{ implies that } [a'^{-1}a'] = e'.$$

Hence by Theorem 7 a' is a scalar and is therefore in Δ'.

Let a in H correspond to a' in H'. Then the complexes $[\Delta a]$ and $[\Delta' a']$ contain elements in one to one correspondence. We have proved this when a is in Δ. Hence suppose a is not in Δ. If $[\delta a]$ is an element of $[\Delta a]$ let a'_1 be the corresponding element in H'. Let δ^{-1} in Δ correspond to δ'^{-1} in Δ'. Then

$$[\delta^{-1}\delta a] = a, \qquad [\delta'^{-1}a'_1] = a', \qquad a'_1 = [\delta' a'],$$

so that a'_1 is in $[\Delta' a']$. We have this result.

THEOREM 24. *If H and H' are simply isomorphic nGM's, and a in H corresponds to a' in H', then the elements of the complexes $[\Delta a]$ and $[\Delta' a']$ are in one to one correspondence, where Δ and Δ' are the nuclii of H and H', respectively. The latter are simply isomorphic groups with respect to the bracket product.*

9. *Automorphism.* If $d = \begin{pmatrix} a \\ a' \end{pmatrix}$ is any substitution on the elements of

an $nG\,H$ which sets up a simple isomorphism of H with itself, we shall call d an *automorphism* of H.

THEOREM 25. *The set of all the automorphisms of a hypergroup forms a group.*

Proof. Clearly the identity substitution $\begin{pmatrix} a \\ a \end{pmatrix}$ is an automorphism; and if $d = \begin{pmatrix} a \\ a' \end{pmatrix}$ is an automorphism, then $d^{-1} = \begin{pmatrix} a' \\ a \end{pmatrix}$ is also an automorphism. Let $d_1 = \begin{pmatrix} a \\ a' \end{pmatrix}$ and $d_2 = \begin{pmatrix} a' \\ a'' \end{pmatrix}$ be two automorphisms. Then $d_1 d_2 = \begin{pmatrix} a \\ a'' \end{pmatrix}$. But since d_1 is an automorphism, the relation

$$a_i a_j = c_1 + c_2 + \cdots + c_n \tag{9}$$

requires that

$$a'_i a'_j = c'_1 + c'_2 + \cdots + c'_n; \tag{10}$$

and since d_2 is an automorphism, the relation (10) requires that

$$a_i'' a_j'' = c_1'' + c_2'' + \cdots + c_n''. \tag{11}$$

Thus the relation (9) implies the relation (11), so that $d_1 d_2$ is an automorphism.

If H is an nGM with nucleus Δ, the substitution

$$\delta' = \begin{pmatrix} a \\ [\delta^{-1} a \delta] \end{pmatrix}, \quad \delta \text{ in } \Delta,$$

is an automorphism of H called an inner automorphism of H.

THEOREM 26. *The set of inner automorphisms of an nGM, H, forms a group which is homomorphic with the nucleus of H. This group is an invariant subgroup of the group of automorphisms of H.*

Proof. Let $\delta'_i = \begin{pmatrix} a \\ [\delta_i^{-1} a \delta_i] \end{pmatrix}$ correspond to δ_i in Δ, and let $\delta_i \delta_j = \delta_k$. Then

$$\delta'_i \delta'_j = \begin{pmatrix} a \\ [\delta_i^{-1} a \delta_i] \end{pmatrix} \begin{pmatrix} a \\ [\delta_j^{-1} a \delta_j] \end{pmatrix}$$

$$= \begin{pmatrix} a \\ [\delta_i^{-1} a \delta_i] \end{pmatrix} \begin{pmatrix} [\delta_i^{-1} a \delta_i] \\ [\delta_j^{-1} \delta_i^{-1} a \delta_i \delta_j] \end{pmatrix} = \begin{pmatrix} a \\ [\delta_k^{-1} a \delta_k] \end{pmatrix} = \delta'_k.$$

Thus the inner automorphisms of H form a group homomorphic with Δ.

Now let

$$\begin{pmatrix} a \\ a' \end{pmatrix}$$

denote any automorphism of H. Then

$$\begin{pmatrix} a \\ a' \end{pmatrix}^{-1}\begin{pmatrix} a \\ [\delta^{-1}a\delta] \end{pmatrix}\begin{pmatrix} a \\ a' \end{pmatrix}=\begin{pmatrix} a \\ a' \end{pmatrix}\begin{pmatrix} a \\ [\delta^{-1}a\delta] \end{pmatrix}\begin{pmatrix} a \\ a' \end{pmatrix}$$
$$=\begin{pmatrix} a' \\ [\delta^{-1}a\delta] \end{pmatrix}\begin{pmatrix} a \\ a' \end{pmatrix}=\begin{pmatrix} a' \\ [\delta^{-1}a\delta] \end{pmatrix}\begin{pmatrix} [\delta^{-1}a\delta] \\ [\delta'^{-1}a'\delta'] \end{pmatrix}=\begin{pmatrix} a' \\ [\delta'^{-1}a'\delta'] \end{pmatrix}.$$

Thus the transform of any inner automorphism of H by any automorphism of H is also an inner automorphism of H. The group of inner automorphisms is an invariant subgroup of the group of automorphisms.

PART III.

Examples of Hypergroups.

10. *Three special examples.* Simple examples of 2-groups are as follows.

Example 1. The integers $1, -1, 2, -2$ if the rule of combination is ordinary multiplication followed by a partition of the product into the sum of two elements of the set. This is a $2G_1M$.

Example 2. The numbers $\cos 0, \cos 1, \cos 2, \cdots$ if the product is defined as twice the ordinary product according to the formula

$$(\cos m)(\operatorname{con} n) = \cos(m+n) + \cos(m-n).$$

This is a $2G_1M$ of infinite order. The identity element is $\cos 0$, the only scalar. The inverse of $\cos n$ is $\cos n$.

Example 3. The system defined by this multiplication table

	e	a	b
e	$e+e$	$a+a$	$b+b$
a	$a+a$	$a+b$	$e+b$
b	$b+b$	$e+b$	$a+a.$

Two other examples were given in § 3.

11. *Double cosets.* Let A be an ordinary group of which S and T are subgroups; and let $(A; S, T)$ denote the set of all the distinct double cosets SaT where a is in A. It is clear that SaT is completely determined by any one of its members. Thus if b is in SaT, then $SaT = SbT$.

We shall define the product of two elements SaT and SbT of $(A; S, T)$ by the formula [4]

[4] If S, T are not both finite this product contains an infinite number of terms. We shall suppose henceforth that S and T are finite groups.

$$(SaT)(SbT) = \Sigma(SatsbT)$$

where the summation runs independently over the elements t of T and s of S. The system so defined is a hypergroup. If 1 is the identity element in A, then $(S1T)$ is an identity of the hypergroup. An inverse of (SaT) is $(Sa^{-1}T)$, where a^{-1} is the inverse of a in A. The associative law holds since it holds in A.

If H is a hypergroup, then we shall define the transpose, H', of H as follows. Let h of H correspond to h' of H' in such a way that when

$$h_1h_2 = a_1 + a_2 + \cdots + a_n, \qquad h_2h_1 = b_1 + b_2 + \cdots + b_n,$$

then

$$h'_1h'_2 = b'_1 + b'_2 + \cdots + b'_n, \qquad h'_2h'_1 = a'_1 + a'_2 + \cdots + a'_n.$$

THEOREM 27. *The hypergroup $(A; T, S)$ is the transpose of the hypergroup $(A; S, T)$.*

Proof. If b is any element of (SaT), then b^{-1} is in $(Ta^{-1}S)$, and if b is in $(Ta^{-1}S)$ then b^{-1} is in (SaT). We let (SaT) of $(A; S, T)$ correspond to $(Ta^{-1}S)$ of $(A; T, S)$. The theorem now follows at once.

If $S = T = 1$, the identity of A, then $(A; S, T)$ reduces to the ordinary group A. If $T = 1$, $S \neq 1$, then $(A; S, T) = (A; S, 1)$ is simply isomorphic with the Mischgruppe of Loewy and Baer, insofar as multiplication is defined for the Mischgruppe, provided we use the bracket product.

Let A be a substitution group on the symbols $1, 2, 3, \cdots, k$. In (SaT) let there be p_{ij} substitutions which carry i into j, and form the matrix (p_{ij}) associated with (SaT). It is easy to prove the following theorem.

THEOREM 28. *If A is a substitution group on the symbols $1, 2, 3, \cdots, k$; and $\alpha = (SaT)$, $\beta = (SbT)$ are two elements of $(A; S, T)$ with product*

$$\alpha\beta = \gamma_1 + \gamma_2 + \cdots + \gamma_n$$

where $\gamma_i = (Sa_iT)$; and $\alpha', \beta', \gamma'_1, \gamma'_2, \cdots, \gamma'_n$ are the matrices associated with $\alpha, \beta, \gamma_1, \gamma_2, \cdots, \gamma_n$, respectively, then

$$\alpha'\beta' = \gamma'_1 + \gamma'_2 + \cdots + \gamma'_n.$$

The matrix associated with (SaT) is the transpose of the matrix associated with $(Ta^{-1}S)$.

Two different double cosets may have one and the same associated matrix. For example, the double cosets

$$(15)(24),\ (1542),\ (14235),\ (2354),\ (13425),\ (13)(254)$$

and

$$(254),\ (1425),\ (15423),\ (15)(234),\ (13542),\ (135)(24)$$

both have the associated matrix

$$\begin{pmatrix} 1 & 0 & 2 & 1 & 2 \\ 1 & 0 & 2 & 1 & 2 \\ 1 & 0 & 2 & 1 & 2 \\ 0 & 6 & 0 & 0 & 0 \\ 3 & 0 & 0 & 3 & 0 \end{pmatrix}.$$

In this example A is the symmetric group of degree 5, $S = \{(123)\}$, $T = \{(14)\}$. The two double cosets are $S(15)(24)T$ and $S(254)T$, respectively. This hypergroup contains no scalars and is of dimension 6, and order 20.

If two elements a and b of a hypergroup H have the property that $ax = bx$ $(xa = xb)$ for every x in H while $xa \neq xb$ $(ax \neq bx)$ for at least one x in H, then a and b are *semi-equal* elements. In the above example $a = S(24)T$ and $b = S(34)T$ are semi-equal: $ax = bx$, $xa \neq xb$ for every x.

In an ordinary group if three elements a, b, x are selected, an element y can be found such that $ax = by$. This is not in general true of a hypergroup. In the above example if $a = S(45)T$, $b = S(12)(45)T$, then for every x there is a y such that $ax = by$. Another exceptional pair of elements having this property is ST and $S(12)T$. For semi-equal elements $y = x$.

12. *Properties of the hypergroup* $(A; S, T)$. Let $B = \{S, T\}$ be the group generated by S and T, and let C be the group of all elements of A which are commutative with B[5]. If $a \in B$ then every element in the double coset SaT is in B. Let $(A; S, T)_0$ be the subset of elements of $(A; S, T)$ which are double cosets made up of elements of B.

THEOREM 29. *$(A; S, T)_0 = (B; S, T)$ is a sub-hypergroup of $(A; S, T)$ which has the following properties. The product of an element in $(A; S, T)_0$ and an element of $(A; S, T)$ not in $(A; S, T)_0$ is a complex of elements not in $(A; S, T)_0$; $(A; S, T)$ can have no identity element outside $(A; S, T)_0$: no element in $(A; S, T)_0$ can have an inverse outside $(A; S, T)_0$.*

Proof. The last two statements follow from the first. Let (SaT) be in $(A; S, T)_0$ and (SbT) not in $(A; S, T)_0$. Then

$$(SaT)(SbT) = \Sigma(SatsbT).$$

[5] C is the "normalizer" of B in A.

Now if $atsb = c$ is in B, then $b = a^{-1}t^{-1}s^{-1}c$ is in B, contrary to hypothesis. Likewise the complex $(SbT)(SaT)$ contains no elements in $(A; S, T)_0$.

If a is in C then every element of (SaT) is in C. Let $(A; S, T)_1$ be the subset of elements of $(A; S, T)$ which are double cosets made up of elements of C. The following theorem is evident.

THEOREM 30. *$(A; S, T)_0$ is a sub-hypergroup of $(A; S, T)_1$; and $(A; S, T)_1 = (C; S, T)$ is a sub-hypergroup of $(A; S, T)$. The product of an element in $(A; S, T)_1$ and an element outside $(A; S, T)_1$ is a complex of elements outside $(A; S, T)_1$; no element in $(A; S, T)_1$ can have an inverse outside $(A; S, T)_1$.*

It will be seen that if U is the group of all elements common to S and T, then each element of $(A; S, T)$ is a sum of elements of $(A; U, U)$; and that each element of $(A; B, B)$ is a sum of elements of $(A; S, T)$.

Let p be any element of A, $S' = p^{-1}Sp$, $T' = p^{-1}Tp$. Let $a' = p^{-1}ap$, where a is in A. Then we have

$$(SaT)(SbT) = \Sigma(SatsbT), \qquad (S'a'T')(S'b'T') = \Sigma(S'[atsb]'T').$$

We therefore have this result.

THEOREM 31. *The hypergroups $(A; S, T)$ and $(A; p^{-1}Sp, p^{-1}Tp)$, where p is in A, are simply isomorphic hypergroups.*

The totality of all the elements of the hypergroups $(A; S_i, T_j)$, $i, j = 1, 2, 3, \cdots$, where S_i, T_i are conjugates of S_1, T_1, respectively, in A, forms a system satisfying postulates I, II § 1, if we define the product of two elements by the formula

$$(S_iaT_j)(S_kbT_l) = \Sigma(S_iat_js_kbT_l),$$

where the summation is taken over the elements t_j of T_j and s_k of S_k.

13. *Properties of the hypergroup $(A; U, U)$.* We have seen that every hypergroup $(A; S, T)$ is a set of complexes of elements of a hypergroup $(A; U, U)$ where U is the set of elements common to S and T.

THEOREM 32. *The hypergroup $(A; U, U)$ is of type M and every element has a unique inverse.*

Proof. The element $(U1U)$ is evidently a scalar, so that $(A; U, U)$ is of type M. If $(U1U) \varepsilon (UaU)(UbU)$ then there exists an element $u' \varepsilon U$ such that

$$au'b = 1, \qquad b = u'^{-1}a^{-1}, \qquad (UbU) = (Ua^{-1}U).$$

Hence every element has a unique inverse.

In the product of two elements of $(A; U, U)$ every element is repeated n times, where n is the order of U; and

$$(UaU)(UbU) = n\Sigma(UaubU)$$

where the summation is taken over the elements u of U. The system $(A; U, U)'$ made up of the elements of $(A; U, U)$ but with the product definition

$$(UaU)(UbU) = \Sigma(UaubU)$$

is a hypergroup of dimension n.

Example. Let U be of order 2. Then $(A; U, U)'$ is a 2-group. Let

$$(UaU)(Ua^{-1}U) = (U1U) + (UcU).$$

Clearly $c^2 = 1$, so that (UcU) is its own inverse. This shows that the hypergroup of Example 3, § 10, cannot be represented as a hypergroup $(A; U, U)'$, where U is of order 2.

The matrix associated (see § 11) with (UaU) is the transpose of the matrix associated with $(Ua^{-1}U)$.

14. *Cyclic hypergroups.* If a hypergroup H is generated by a single element a of H, then H will be called a *cyclic hypergroup.* In a cyclic hypergroup the commutative law does not necessarily hold. We may write

$$aa = a^2, aaa = a^3, \cdots;$$

and clearly $a^m a^n = a^{m+n}$. It is evident that the set of elements generated by a single element of a hypergroup is closed, i. e. satisfies postulate I, § 1, but I am unable to say whether or not it forms a hypergroup.

In connection with a cyclic nG one would naturally seek an analog of the representation of a cyclic group as the group of rotations of a regular polygon. This leads to the notion of cyclomorphism. Let H be any nG, and let $d = \binom{a}{a'}$

be a substitution on the elements of H such that $ka = \Sigma c_i$ implies that $ka' = \Sigma c'_i$ where $a \sim a'$ and $c_i \sim c'_i$ $(i = 1, 2, 3, \cdots, n)$, for every a in H, and k is an assigned element of H. Then d will be called a *cyclomorphism* of H relative to k. The set $G(k)$ of all cyclomorphisms of H relative to k is clearly a group.

If H is cyclic with generator k we may represent it in a certain sense as a system of points and directed lines. Let each element of H be identified with a certain point in a space S. From the point k draw directed lines to each of the points in the set $[k^2]$ which are different from k. Call them $k_1, k_2, \cdots, k_p$. If k_i occurs m_i times in k^2 we shall consider the line from k to k_i as an m_i-tuple line. Then from each point k_i draw lines to the elements in $[kk_i]$ which are different from k_i, etc. The group $G(k)$ of cyclomorphisms of H relative to k is a group of transformations of this configuration of points and lines into itself.

As an illustration let $H = (A; U, U)'$ where A is the symmetric group of degree 5, and $U = \{(12)\}$. Then H is of order 33 with the elements

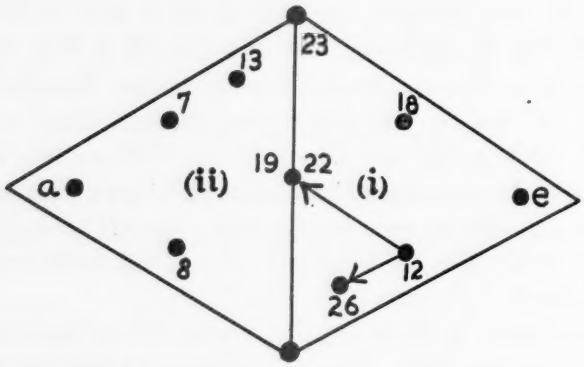
$$e = U1U, \quad a = U(34)U, \quad b = U(35)U, \quad c = U(45)U, \quad d = U(345)U,$$
$$f = U(354)U$$

which are scalars; and the non-scalars

$$1 = U(13)U, \quad 2 = U(14)U, \quad 3 = U(15)U, \quad 4 = U(134)U, \quad 5 = U(135)U,$$
$$6 = U(143)U, 7 = U(145)U, 8 = U(153)U, 9 = U(154)U, 10 = U(1324)U,$$
$$11 = U(1325)U, \quad 12 = U(1345)U, \quad 13 = U(1354)U, \quad 14 = U(1425)U,$$
$$15 = U(1435)U, \quad 16 = U(1453)U, \quad 17 = U(1534)U, \quad 18 = U(1543)U,$$
$$19 = U(13245)U, \quad 20 = U(13254)U, \quad 21 = U(13425)U, \quad 22 = U(13524)U,$$
$$23 = U(14253)U, 24 = U(14325)U, 25 = U(13)(45)U, 26 = U(14)(35)U,$$
$$27 = U(15)(34)U.$$

This $2G_1M$ may be generated by $k = 12$; and $G(k)$ is simply isomorphic with the group of rotations of a six sided double pyramid into itself. This suggests that we represent the elements as points upon this double pyramid.

Hold the solid with the principal axis vertical and, looking downward, number the faces of the upper tetrahedron i, iv, iii counterclockwise. Number the faces below these on the lower tetrahedron ii, v, vi, respectively. On face i mark the elements 23, 18, *e*, 12, 26, 22 as shown in the accompanying figure.



The faces iv and iii have these elements replaced by 20, 1, *d*, 15, 9, 14 and 10, 16, *f*, 3, 4, 21, respectively. The face numbered ii is marked with the elements 8, *a*, 7, 13, 19 as shown. The faces v and vi have these elements replaced by 25, *b*, 27, 2, 24 and 6, *c*, 5, 17, 11, respectively. The points representing 19 and 22, 14 and 24, 11 and 21 are considered as double points. Now $12 \cdot 12 = 26 + 22$. Draw arrows from 12 to 26 and to 22. Then $12 \cdot 26 = 18 + 24$. Draw arrows from 26 to 18 and to 24. Likewise draw arrows from 22 to 23 and to 20, and so on until the configuration is complete. The six rotations of the solid into itself carry the configuration of points and lines into itself.

Added in proof. Since this paper was written a paper by M. F. Marty: "Sur les groupes et hypergroupes," *Annales de l'École Normale,* (3), vol. 53, pp. 83-123, has appeared. Marty applies the special hypergroup $[A; U, U]$ of § 13 to a problem in topology. His results do not overlap mine.

NORTHWESTERN UNIVERSITY,
EVANSTON, ILL.

PRIMARY ABELIAN GROUPS AND THEIR AUTOMORPHISMS.[1]

By REINHOLD BAER.

The object of this paper is to investigate the relations between automorphisms and properties of the group itself. To every group of automorphisms corresponds first the group of its fixed elements and secondly the subgroup, generated by the "commutators" $x - xf$ of the elements x of the group under the automorphisms f of the given group of automorphisms. This commutator group of a group of automorphisms A is exactly the smallest subgroup S such that A induces mod S the identical automorphism. Those subgroups of primary Abelian groups which are complete groups of fixed elements and those which are complete commutator groups of suitable groups of automorphisms are characterized by their structure and their situation in the whole group.

Furthermore those automorphisms are studied which map every subgroup upon itself. It is proved that the centralizer of the group of all the automorphisms is essentially formed by these automorphisms.

Finally some instances of the problem to characterize a group by properties of its automorphisms are considered. In this respect the following two classes of primary Abelian groups G are enumerated:

The group of all the automorphisms of G is Abelian.

Every automorphism of the lattice of the subgroups of G is induced by an automorphism of G.

The proofs are based on certain decomposition-theorems for primary Abelian groups which are enunciated and proved in section 1.

1. Invariants and decompositions of primary Abelian groups.[2]

If G is an (additively written primary Abelian) group (of characteristic p), then the *order* of every element x in G is a power $p^{n(x)}$ of the prime number p. The elements x in G with $n(x) \leq i$ form a characteristic subgroup $(G, n(x) \leq i)$ of G.

The *rank* $r(G)$ of the group G is the number of elements in $(G, n(x) \leq 1)$, if $(G, n(x) \leq 1)$ is infinite, and if $(G, n(x) \leq 1)$ is finite, then $(G, n(x) \leq 1)$

[1] Presented to the American Mathematical Society, October 31, 1936. Received by the Editors September 22, 1936.

[2] In this section, those concepts and facts of the theory of primary Abelian groups are collected which are needed in the course of this paper.

contains exactly $p^{r(G)}$ elements. Thus G contains $r(G)$ elements, if $r(G)$ is infinite.

Since the orders of the elements of $(G, n(x) \leq i)$ are bounded, every $(G, n(x) \leq i)$ with positive i is a direct sum of $r(G)$ cyclic groups.[3]

Denote by $p^{\omega}G$ the intersection of all the groups p^iG (for integral i)[4] and by $p^{\infty}G$ the uniquely determined greatest subgroup S of G, satisfying $S = pS$. Every subgroup S of G which satisfies $S = pS$ is a direct summand of G and a direct sum of groups of type p^{∞}.[5,6]

Put $r(G, i) = r((p^{i-1}G, n(x) \leq 1)/(p^iG, n(x) \leq 1))$ for integral i and $r(G, \infty) = r(p^{\infty}G)$. Then the structure of G is completely determined by these invariants, provided G is a direct sum of cyclic groups and of groups of type p^{∞}. If in particular G is the direct sum of groups G_v, i. e.

$$G = \sum_v G_v,$$

then at most $r(G, i)$ of the groups G_v are cyclic groups of order p^i and at most $r(G, \infty)$ of the groups G_v are groups of type p^{∞}.

Conversely: If m is a positive integer, then

$$G = \sum_{i=1}^{m} G_i + p^{\infty}G + G',$$

where G_i is a direct sum of $r(G, i)$ cyclic groups of order p^i, $p^{\infty}G$ a direct sum of $r(G, \infty)$ groups of type p^{∞} and

$$r(G', 1) = \cdots = r(G', m) = r(G', \infty) = 0.$$

A proof of this theorem[7] may be given: $G = G'' + p^{\infty}G$, as has been mentioned before, and $r(G'', i) = r(G, i)$ for every finite i, $r(G'', \infty) = 0$. There exists for every integer i a subset B_i of G'' such that the elements $p^{i-1}b$ with b in B_i form a basis of $(p^{i-1}G'', n(x) \leq 1) \bmod (p^iG'', n(x) \leq 1)$ (and of $(p^{i-1}G, n(x) \leq 1) \bmod (p^iG, n(x) \leq 1)$). The subgroup G_i of G'', generated by B_i, is then a direct sum of $r(G, i)$ cyclic groups of order p^i.

Denote by M the subgroup of G'', generated by the elements in sets B_i

---

[3] A proof of this well known theorem may be found e. g. in R. Baer, "Der Kern, eine characteristische Untergruppe," *Comp. Math.*, vol. 1 (1934), pp. 254-283, Lemma in § 5.

[4] p^iG consists of all the elements p^ix for x in G and it is obvious how to define p^vG for every ordinal number v by complete transfinite induction.

[5] A group of type p^{∞} is generated by elements g_i, satisfying $n(g_0) = 1$, $pg_i = g_{i-1}$.

[6] Cf. e. g. L. Zippin, "Countable torsion groups," *Annals of Mathematics*, vol. 36 (1935), pp. 86-99.

[7] Similar results and arguments may be found in the investigations of H. Prüfer.

with $m < i$, and let G' be the set of all elements x in G'' such that there exists an element x' in M, satisfying $x \equiv x' \bmod pG''$. Then G' is a subgroup of G'', G is the direct sum of the groups $G_1, \cdots, G_m, p^\infty G, G'$ and this direct decomposition of G meets the requirements of the above proposition.

It is an important corollary, derived from the above proof, that the set B_i is a basis of G_i and may be chosen at random, only restricted by the condition that the elements $p^{i-1}b$ with b in B_i form a basis of

$$(p^{i-1}G, n(x) \leq 1) \bmod (p^i G, n(x) \leq 1).$$

If N is the subgroup of G, generated by the elements in all the sets B_i of the above construction, then $G/N = p(G/N)$. If the orders of the elements in G are not bounded, then it is possible to choose the sets B_i in such a way that $G/N \neq 0$. This implies the following theorem:

G is homomorphic to a group of type p^∞ if, and only if, the orders of the elements in G are not bounded.[8]

2. Automorphisms.

Every single-valued additive function $\boldsymbol{f}$ of the elements of the (additively written Abelian) group G with values $x\boldsymbol{f}$ in G is an automorphism of G. The set $\boldsymbol{R}(G)$ of all the automorphisms of G is a ring under the well known definition of addition and multiplication of automorphisms.

The proper automorphisms of G (which define a one-one correspondence between all the elements of G such that an inverse automorphism exists) form a multiplicative group $\boldsymbol{P}(G)$. The identity element of $\boldsymbol{P}(G)$ may be denoted by **1**.

If $\boldsymbol{S}$ is any set of automorphisms of G, $g(x) = \sum_{i=0}^{n} c_i x^i$ a polynomial with integer coefficients, S and T subsets of G, then

$(S, xg(\boldsymbol{S}) \leq T)$ is the set of all the elements x in S which are mapped by all the automorphism $g(\boldsymbol{s})$ with $\boldsymbol{s}$ in $\boldsymbol{S}$ upon elements in T; and

$(\boldsymbol{S}, Sg(\boldsymbol{x}) \leq T)$ is the set of all the automorphisms $\boldsymbol{x}$ in $\boldsymbol{S}$ such that $g(\boldsymbol{x})$ maps S upon a subset of T.[9]

Some particular instances of the correspondence [10] between sets of auto-

[8] Cf. R. Baer, "Dualisms in Abelian groups," *Bulletin of the American Mathematical Society.*

[9] Also $(S, T \leq xg(\boldsymbol{S}))$ and $(\boldsymbol{S}, T \leq Sg(\boldsymbol{x}))$ may be considered.

[10] Such a correspondence has first been introduced by K. Shoda, "Über den Automorphismenring bezw. die Automorphismengruppe einer endlichen Abelschen Gruppe," *Proceedings of the Imperial Academy of Japan*, vol. 6 (1930), p. 10.

morphisms and sets of elements, thus defined, shall be investigated in the course of this paper.

It is easily verified, that

to every subgroup S of G and to every isomorphism $\mathfrak{g}$ of G/S upon a subgroup T of G there exists one and only one automorphism $\mathfrak{f}$ of G such that $S = (G, x\mathfrak{f} = 0)$ and $\mathfrak{g}$ is the isomorphism of G/S upon $T = G\mathfrak{f}$, induced by $\mathfrak{f}$. Conversely every automorphism $\mathfrak{f}$ induces an isomorphism of $G/(G, x\mathfrak{f} = 0)$ upon $G\mathfrak{f}$.

Since the automorphisms $\mathfrak{f}$ and $\mathfrak{1} - \mathfrak{f}$ determine the one the other completely, the above result implies, that the automorphism $\mathfrak{f}$ of G is completely determined by the group $(G, x\mathfrak{f} = x)$ of the fixed elements and the induced isomorphism of $G/(G, x\mathfrak{f} = x)$ upon $G(\mathfrak{1} - \mathfrak{f})$.

3. Fixed elements.

NOTATION. If S is a subgroup of G, then $F(S)$ is the subgroup of all those elements of G which are mapped upon themselves by all the proper automorphisms of G which map every element of S upon itself, i. e.

$$F(S) = (G, x(\mathfrak{P}(G), S(\mathfrak{1} - \mathfrak{f}) = 0) = x).$$

Then $S \leq F(S)$ for every S.

THEOREM 3. 1.[11] *Suppose that G is a primary Abelian group of characteristic p, and that $p^\omega G = p^\infty G$. Then*

$$F(0) \neq 0$$

if, and only if,

(a) $p = 2$;

(b) *either* $r(G, \infty) = 1$ *or there exists an integer m such that*

$$r(G, m) = 1, \qquad p^m G = 0.$$

If the conditions (a) *and* (b) *are satisfied, then $F(0)$ is a group of order 2 and*

$$F(0) = (p^\infty G, n(x) \leq 1) \quad \text{or} \quad F(0) = (p^{m-1}G, n(x) \leq 1) = p^{m-1}G.$$

Proof. Since $x\mathfrak{r} = -x$ is a proper automorphism of G, $2x = 0$ for every element x in $F(0)$. Thus $F(0) \neq 0$ implies $p = 2$.

If g and h are different elements of order 2 in G, then

$$G = A + B + G',$$

[11] Cf. R. Baer, "Types of elements and characteristic subgroups of Abelian groups," *Proceedings of the London Mathematical Society*, ser. 2, vol. 39 (1935), pp. 481-514, in particular Theorem 10.

where g is the only element of order 2 in A, h the only element of order 2 in B, as has been pointed out in section 1. A and B are either cyclic groups or of type 2^∞ and it may therefore be assumed that A is isomorphic with a subgroup of B. Then there exists an automorphism of G which maps every element of $B + G'$ upon itself and g upon $g + h$, i. e. g does not belong to $F(0)$. The necessity of the conditions is now a consequence of the theorems in section 1.

The conditions are sufficient, since a group of order 2 contains but one element $\neq 0$, and since $(p^iG, n(x) \leq 1)$ and $(p^\infty G, n(x) \leq 1)$ is mapped upon itself by every proper automorphism of G.

LEMMA 3. 2. *Assume that S is a subgroup of the primary Abelian group G of characteristic p.*

(a) *If G/S is a cyclic group, then $F(S)$ is the join of S and $F(0)$.*

(b) *If G/S is a group of type p^∞, then*

$$F(S) = G \qquad \text{for } p^\infty G = 0,$$
$$= \text{join of } S \text{ and } F(0) \text{ for } p^\infty G \neq 0.$$

Proof. Note first that $S \leq F(S)$ and $F(0) \leq F(S)$.

If G/S is a cyclic group of order p^n with $0 < n$, then there exists an element g in G which generates G mod S. Put $g^* = p^n g$ and $m = n(g^*)$. Every element of G can be represented in one and only one way in the form

$$s + cg \text{ with } s \text{ in } S \text{ and } 0 \leq c < p^n.$$

If not at the same time $m = 0$ and $p = 2$, i. e. if $1 + p^m$ is relatively prime to p, then a proper automorphism $\mathfrak{f}$ of G is defined by

$$(s + cg)\mathfrak{f} = s + c(1 + p^m)g$$

and

$$S = (G, x\mathfrak{f} = x).$$

If $m = 0$ and $p = 2$, then $g^* = 0$ and $G = \bar{g} + S$ where $\bar{g}$ is the cyclic group, generated by g, since $\bar{g}$ and S generate G and have only 0 in common. If S contains an element h of order 2^n, then an automorphism $\mathfrak{f}$ of G is defined by

$$g\mathfrak{f} = g + h, \quad s\mathfrak{f} = s \text{ for } s \text{ in } S$$

and

$$S = (G, x\mathfrak{f} = x).$$

If S does not contain an element of order 2^n, then $r(G, n) = 1$, $2^n G = 0$, i. e. $F(0)$ is generated by $2^{n-1}g$. An automorphism $\mathfrak{f}$ of G is defined by

$$gf = -g, \quad xf = x \text{ for } x \text{ in } S$$

and

$$(G, xf = x) = \text{join of } S \text{ and } F(0).$$

Assume now that G/S is a group of type p^∞. If f is an automorphism of G such that $sf = s$ for every s in S, then $1 - f$ maps G and G/S upon a subgroup of G of type p^∞ or upon 0. Thus $p^\infty G = 0$ implies $G = F(S)$.

Assume now that $p^\infty G \neq 0$.

If there exists a subgroup Z of type p^∞ of G which contains elements $\neq 0$ of S, then denote by $z(i)$ a sequence of elements in Z such that $pz(i) = z(i-1)$ and $n(z(1)) = 1$. $z(1)$ is an element of S. There exists furthermore a sequence $g(i)$ in G such that $pg(i) \equiv g(i-1) \bmod S$ and $n(S + g(1)) = 1$. A proper automorphism f of G is defined by

$$sf = s \text{ for } s \text{ in } S, \qquad g(i)f = g(i) + z(i),$$

since $g(1)f \equiv g(1) \bmod S$,

$$pg(i)f = pg(i) + pz(i) \equiv g(i-1) + z(i-1) \equiv g(i-1)f \bmod S,$$

and since G/S is a group of type p^∞. Since $g(i)f \neq g(i)$, $S = (G, xf = x)$.

If none of the subgroups of type p^∞ of G contains an element $\neq 0$ of S, then $p^\infty G$ is a group of type p^∞, since G/S is of type p^∞, and consequently

$$G = S + p^\infty G.$$

Since by Theorem 3.1, $F(0) = (p^\infty G, n(x) \leq 1)$, the automorphism f, defined by $sf = s$ for s in S, $xf = -x$ for x in $p^\infty G$, satisfies

$$(G, xf = x) = \text{join of } S \text{ and } F(0).$$

Thus the proof of the Lemma 3.2 is complete. Note that in the discussed cases $F(S)$ is the group of fixed elements of a single proper automorphism.

THEOREM 3.3. *Assume that S is a subgroup of the primary Abelian group G of characteristic p.*

(a) *If $p^\infty G \neq 0$, then $F(S)$ is the join of S and $F(0)$.*

(b) *If $p^\omega G = 0$, then $F(S)$ is the join of $W(S)$ and $F(0)$ where $W(S)$ is the subgroup of G, satisfying $S \leq W(S)$ and $W(S)/S = p^\omega(G/S)$.*[12]

Proof. If f is any proper automorphism of G such that the elements of

[12] Since every characteristic subgroup $\neq 0$ of a direct sum of irreducible groups has the form, required in (a) and (b) respectively, Theorem 11 of the paper, mentioned in footnote [11] is a consequence of this Theorem.

S are fixed elements of $\boldsymbol{f}$, then $\mathbf{1} - \boldsymbol{f}$ maps G/S homomorphic upon a subgroup of G and maps therefore $W(S)/S = p^\omega(G/S)$ upon a subgroup of $p^\omega G$. Thus $W(S) \leq F(S)$, if $p^\omega G = 0$. Since $p^\omega(G/W(S)) = 0$, it follows from the theorems of the section 1, that every element $W(S) + g$ of order p is contained in a cyclic direct summand of $G/W(S)$. There exists therefore for every element g of G which is not contained in $W(S)$ a subgroup H of G such that $W(S) \leq H$, G/H is a cyclic group and g is not contained in H. Then it follows from Lemma 3. 2, (a) that $F(H)$ is the join of $F(0)$ and H and now it is easily seen that $F(S)$ is the join of $F(0)$ and $W(S)$, if $p^\omega G = 0$.

If $p^\infty G \neq 0$, then denote by S' the join of S and $F(0)$. If g is any element of G which is not contained in S', then there exists a greatest subgroup H of G, containing S' but not g. G/H is either a cyclic group or a group of type p^∞ and it follows from Lemma 3. 2 that $H = F(H)$, since $F(0) \leq H$ and this completes the proof.

Since (as pointed out in section 1) G is not homomorphic to a group of type p^∞ if, and only if, the orders of the elements of G are bounded, Lemma 3. 2 and Theorem 3. 3 imply the

COROLLARY 3. 4. *$F(S)$ is for every subgroup S of G the join of S and $F(0)$ if, and only if, either $p^\infty G \neq 0$ or the orders of the elements of G are bounded.*

COROLLARY 3. 5. *The subgroup S of the primary Abelian group G with $p^\omega G = 0$ satisfies $F(S) = S$ if, and only if, $F(0) \leq S$ and $p^\omega(G/S) = 0$.*

4. Commutators.

NOTATIONS. If $\boldsymbol{S}$ is any group of (proper) automorphisms of the Abelian group G, then $(G, \boldsymbol{S})$ is the subgroup of G, generated by the elements $x - x\boldsymbol{s}$ with x in G and $\boldsymbol{s}$ in $\boldsymbol{S}$.

If S is any subgroup of G, then $(\boldsymbol{P}(G), S)$ is the set (and therefore the group) of all the proper automorphisms $\boldsymbol{f}$ of G such that $x - x\boldsymbol{f}$ is an element of S for every x in G.

$C(S) = (G, (\boldsymbol{P}(G), S))$ is a subgroup of G, contained in S.

$(G, \boldsymbol{S})$ is the cross cut of all the subgroups T of G such that $\boldsymbol{S}$ induces the identical automorphism in G/T.

$(G, \boldsymbol{S}) \leq (G, \boldsymbol{T})$, if $\boldsymbol{S} \leq \boldsymbol{T}$ and $(\boldsymbol{P}(G), S) \leq (\boldsymbol{P}(G), T)$, if $S \leq T$.

$(G, \boldsymbol{S}) = (G, (\boldsymbol{P}(G), (G, \boldsymbol{S})))$, $(\boldsymbol{P}(G), S) = (\boldsymbol{P}(G), (G, (\boldsymbol{P}(G), S)))$.

$(\boldsymbol{P}(G), G) = \boldsymbol{P}(G)$, $(\boldsymbol{P}(G), 0) = \mathbf{1}$.

$(G, \mathbf{1}) = 0$.

If $\boldsymbol{S}$ is a self conjugate subgroup of $\boldsymbol{P}(G)$, then $(G, \boldsymbol{S})$ is a characteristic subgroup of G, and if S is a characteristic subgroup of G, then $(\boldsymbol{P}(G), S)$ is a self conjugate subgroup of $\boldsymbol{P}(G)$.

THEOREM. *Assume that S is a subgroup of the primary Abelian group G of characteristic p.*

(a) *If the orders of the elements of $G/p^\omega G$ are not bounded, then $S = C(S)$.*

(b) *If p^m is the maximum order of the elements of $G/p^\omega G$ and either $p \neq 2$ or $p = 2$, but $1 < r(G, m)$, then $C(S)$ is the join of $p^\infty S$ and $(S, n(x) \leq m)$.*

(c) *If $p = 2$, 2^m is the maximum order of the elements of $G/2^\omega G$ and $r(G, m) = 1$, then $C(S)$ is the join of $2^\infty S$, $(S, n(x) < m)$ and $(S; n(x) \leq m, 2^{m-1}x \text{ in } 2^\omega G)$.*

Proof. Note first that $C(S) \leq S$ for every S.

If i is an integer such that $r(G, i) \neq 0$, then there exists a cyclic direct summand of order p^i in G, i. e.

$$G = Z + Z'$$

where Z is a cyclic group of order p^i, generated by z. If s is any element in S such that $n(s) < i$, then an automorphism $\boldsymbol{f}$ in $(\boldsymbol{P}(G), S)$ is defined by

$$x\boldsymbol{f} = x \text{ for } x \text{ in } Z', \qquad z\boldsymbol{f} = z + s$$

and s is an element of $C(S)$. Hence

$$(S, n(x) < i) \leq C(S) \text{ for every } i \text{ with } r(G, i) \neq 0.$$

If the orders of the elements of $G/p^\omega G$ are not bounded, then there exists an infinity of integers i such that $r(G, i) \neq 0$, since $r(G, i) = r(G/p^\omega G, i)$ for every finite i. This proves (a), since S is the join of all the groups $(S, n(x) < i)$ with integral i.

If the orders of the elements of $G/p^\omega G$ are bounded, then there exists a maximum order p^m of the orders of the elements of $G/p^\omega G$ and $r(G, m) \neq 0$, $r(G, i) = 0$ for $m < i$. Then the above result shows that

$$(S, n(x) < m) \leq C(S).$$

If $s \neq 0$ is an element of $p^\infty S$, then there exist elements $s(i)$ in $p^\infty S$ such that $s(0) = s$, $ps(i) = s(i-1)$ for $0 < i$. The elements $s(i)$ generate a direct summand D of S and G, i. e.

$$G = D + D'$$

and an automorphism $\boldsymbol{f}$ in $(\boldsymbol{P}(G), S)$ is defined by

$$x\boldsymbol{f} = x \text{ for } x \text{ in } D', \qquad s(i)\boldsymbol{f} = s(i) + s(i-1) \text{ for } 0 < i,$$

since $n(s(i-1)) < n(s(i))$ and

$$ps(i)\boldsymbol{f} = ps(i) + ps(i-1) = s(i-1) + s(i-2) = s(i-1)\boldsymbol{f} \text{ for } 1 < i.$$

s is therefore an element of $C(S)$, i. e.

$$p^{\infty}S \leq C(S).$$

By section 1 there exists a direct decomposition

$$(*) \qquad G = \sum_{i=1}^{m} G_i + p^{\infty}G$$

of G where G_i is a direct sum of $r(G, i)$ cyclic groups of order p^i. If s is an element of order p^m in S, then

$$s = u + v + w$$

with u in $\sum_{i=1}^{m-1} G_i$, v in G_m and w in $p^{\infty}G$, $n(w) \leq m$.

If $n(v) < m$, i. e. if $p^{m-1}s$ is an element of $p^{\infty}G$, then let Z be a cyclic direct summand of order p^m of G, generated by z, i. e.

$$G = Z + Z'$$

and an automorphism $\boldsymbol{f}$ in $(\boldsymbol{P}(G), S)$ is defined by

$$x\boldsymbol{f} = x \text{ for } x \text{ in } Z', \qquad z\boldsymbol{f} = z + s.$$

Consequently s is an element of $C(S)$, i. e.

$$(S; n(x) \leq m, p^{m-1}x \text{ in } p^{\omega}G) \leq C(S).$$

If (in the above representation of the element s in S) v is an element of order p^m, then v generates a direct summand $\bar{v}$ of G, i. e.

$$G = \bar{v} + V$$

and, if $p \neq 2$, an automorphism $\boldsymbol{f}$ in $(\boldsymbol{P}(G), S)$ is defined by

$$x\boldsymbol{f} = x \text{ for } x \text{ in } V, \qquad v\boldsymbol{f} = v + s = 2v + u + w.$$

Consequently s is an element of $C(S)$ and

$$(S, n(x) \leq m) \leq C(S), \text{ if } p \neq 2.$$

If $1 < r(G, m)$ and $n(v) = m$, then

$$G = \bar{v} + \bar{z} + G'$$

where $\bar{z}$ is the cyclic group of order p^m, generated by z. An automorphism $\boldsymbol{f}$ in $(\boldsymbol{P}(G), S)$ is then defined by

$$x\boldsymbol{f} = x \text{ for } x \text{ in } G' + \bar{v}, \qquad z\boldsymbol{f} = z + s,$$

since $p^{m-1}s$ is an element of $\bar{v} + G'$. Thus s is an element of $C(S)$ and

$$(S, n(x) \leq m) \leq C(S), \quad \text{if} \quad 1 < r(G, m).$$

Let finally $\boldsymbol{f}$ be any automorphism in $(\boldsymbol{P}(G), S)$. Then $\mathbf{1} - \boldsymbol{f}$ maps $\sum_{i=1}^{m-1} G_i$ (cf. (*)!) upon a subgroup of $(S, n(x) < m)$, $p^\infty G$ upon a subgroup of $p^\infty S$ and G_m upon a subgroup of $(S, n(x) \leq m)$ and that proves (b). If $p = 2$ and $r(G, m) = 1$, i. e. if G_m is a cyclic group of order 2^m, generated by g, then

$$g\boldsymbol{f} - g = b + cg + d$$

with b in $\sum_{i=1}^{m-1} G_i \leq (G, n(x) < m)$ and d in $p^\infty G$. Since $\boldsymbol{f}$ is a proper automorphism of G, $1 + c$ is relatively prime to $p = 2$, i. e. c is even and $2^{m-1}(g\boldsymbol{f} - g) = 2^{m-1}d$ is an element of $2^\infty G$, i. e. G_m is mapped by $\mathbf{1} - \boldsymbol{f}$ upon a subgroup of $(S; n(x) \leq m, 2^{m-1}x \text{ in } 2^\omega G)$ and this completes the proof of (c).

If in particular $S = G$, then the theorem implies the

COROLLARY. *Assume that G is a primary Abelian group of characteristic p. Then $C(G) \neq G$ if, and only if,*

(1) $p = 2$;

(2) *there exists an integer m such that $r(G, m) = 1$, $2^m G = 2^\omega G$.*

If (1) and (2) are satisfied, then $C(G)$ is the join of $2^\omega G$ and $(G, n(x) < m)$, and $G/C(G)$ is of order 2.

5. Automorphisms of the lattice of the subgroups.

If $L(G)$ is the set of all the subgroups of the Abelian group G, then an

automorphism $\mathfrak{f}$ of $L(G)$ is a single valued function of the elements of $L(G)$ with values in $L(G)$, satisfying:

$$L(G)\mathfrak{f} = L(G),$$

$$S \leq T \text{ if, and only if, } S\mathfrak{f} \leq T\mathfrak{f},$$

S and $S\mathfrak{f}$ contain the same number of elements,
if $S < T$, then T/S and $(T\mathfrak{f})/(S\mathfrak{f})$ contain the same number of elements.[13]

Every automorphism of the lattice $L(G)$ of the subgroups of G is therefore a one-one correspondence between the subgroups of L, mapping 0 upon 0, G upon G.

Every automorphism of G induces an automorphism of $L(G)$, but the converse is generally not true, as the following theorem shows.

THEOREM 5.1. *Assume that G is a primary Abelian group of characteristic p. Then every automorphism of $L(G)$ is induced by an automorphism of G if, and only if, either G is irreducible, or G is a direct sum of a group of order 2 and of an irreducible group.*[14]

Proof. Since by section 1 irreducible primary groups are cyclic groups or groups of type p^∞, it is well known [15] that every automorphism of $L(G)$ is induced by an automorphism of G, if G is irreducible. If $G = T + W$ where T is of order 2 and W irreducible, then there exists to every cyclic subgroup Z of W such that $2W \neq 0$ exactly one subgroup Z' of G such that $Z' \neq Z$, $2Z' = 2Z$. If W is a cyclic group of order 2, then every permutation of the three subgroups of order 2 can be realized by an automorphism of G. If W is a cyclic group of order 2^i with $1 < i$, w an element of order 2^i in W, $z \neq 0$ an element of T, then all the automorphisms of G have the form $z\mathfrak{f} = z + z'$ with $n(z') \leq 1$, z' in W, $w\mathfrak{f} = cw + ez$ where c is an odd number and $e = 0$ or $= 1$, and therefore every automorphism of $L(G)$ is induced by an automorphism of G. If finally W is of type 2^∞, then every automorphism of G maps W upon itself and permutes the elements of order 2 in G which are not contained in W and therefore every automorphism of $L(G)$ is induced by an automorphism of G.

---

[13] The automorphisms of $L(G)$ preserve the situation of the subgroups of G (= situationstreue Abbildung von G); cf. R. Baer, "Situation der Untergruppen und Struktur der Gruppe," *Sitzungs-Berichte Heidelberg Akademie Wissenschaften, Math.-Nat. Kl.*, (2) (1933), pp. 12-17.

[14] The same property is possessed by, e.g., the subgroups of the additive group of the rational numbers.

[15] Cf. the paper, mentioned in footnote [13].

Assume now that $G = Z' + Z''$ where Z' and Z'' are cyclic groups, generated by elements z' and z'' respectively, neither of them of order 2. Then there exists an automorphism of $L(G)$ which

maps the subgroup generated by	upon the subgroup generated by
z'	$z' + p^{n(z'')-1}z''$
$z' + p^{n(z'')-1}z''$	$z' + p^{n(z'')-1}z'' + p^{n(z')-1}z'$
$z' + p^{n(z'')-1}z'' + p^{n(z')-1}z'$	z'

and all the other cyclic subgroups of G upon themselves. Such an automorphism of $L(G)$ cannot be induced by an automorphism of G.

A similar argument shows the existence of automorphisms of $L(G)$ which are not induced by automorphisms of $G = Z' + Z''$, if Z' and Z'' are both irreducible groups $\neq 0$ and not of order 2.

If finally $G = G' + G''$, then every automorphism of $L(G')$ is induced by an automorphism of $L(G)$ and therefore it follows from section 1 and the above considerations that the conditions of the theorem are necessary.

THEOREM 5.2. *The automorphism $\mathfrak{f}$ of the primary Abelian group G of characteristic p induces the identical automorphism in $L(G)$ if, and only if, there exists an integral p-adic number $c \not\equiv 0 \bmod p$ such that $x\mathfrak{f} = cx$ for every element x in G.*[16]

It suffices to prove the necessity of the condition. If the automorphism $\mathfrak{f}$ of G induces in $L(G)$ the identical automorphism, then $\mathfrak{f}$ maps every cyclic subgroup of G upon itself and therefore

$$x\mathfrak{f} = c(x)x \text{ for every } x \text{ in } G.$$

Since $\mathfrak{f}$ is an automorphism of G, the function $c(x)$ satisfies:

$$(c(x+y) - c(x))x = (c(y) - c(x+y))y.$$

If the cyclic groups, generated by x and y have no element $\neq 0$ in common, then

$$c(x+y) \equiv c(x) \bmod p^{n(x)}, \qquad c(x+y) \equiv c(y) \bmod p^{n(y)}$$

and consequently

$$c(x) \equiv c(y) \bmod p^{\text{minimum}(n(x),n(y))}.$$

---

[16] If $\sum_{i=0}^{\infty} c_i p^i$ is an integral p-adic number c, then $\sum_{i=0}^{j} c_i p^i x = \sum_{i=0}^{j+k} c_i p^i x$, if $n(x) \leq j \leq j+k$, and cx is therefore uniquely determined.

Since by section 1 $(G, n(x) \leq i)$ is a direct sum of cyclic groups, and since $\boldsymbol{f}$ induces an automorphism in $(G, n(x) \leq i)$, there exists by the above argument to every integer i an integer c_i such that

$$x\boldsymbol{f} = c_i x \text{ for } n(x) \leq i$$

and this completes the proof, since

$$c_i \equiv c_{i-1} \bmod p^{i-1}.$$

6. The centralizer of the group of automorphisms.

NOTATIONS. $\boldsymbol{Z}(G)$ is the centralizer of the group $\boldsymbol{P}(G)$ of all the proper automorphisms of G, i. e. the group of all those proper automorphisms of G which are permutable with every proper automorphism of G.

$\boldsymbol{Z}^*(G)$ is the group of all those proper automorphisms $\boldsymbol{f}$ of G, satisfying:

$$S\boldsymbol{f} = S \text{ for every subgroup } S \text{ of } G.$$

It is a consequence of Theorem 5. 2 that

$$\boldsymbol{Z}^*(G) \leq \boldsymbol{Z}(G) \text{ for every primary Abelian group } G.$$

THEOREM. *Assume that G is a primary Abelian group of characteristic p.*

(I) $\boldsymbol{Z}^*(G) \neq \boldsymbol{Z}(G)$ *if, and only if,*

(Ia) $p = 2$;

(Ib) *there exists a (uniquely determined) integer $m = m(G)$ such that*

$$1 = r(G, \infty) = r(G, m), \qquad 0 = r(G, k) \text{ for } m < k.$$

(II) *Assume that G satisfies* (Ia) *and* (Ib).

(IIa) *The proper automorphism $\boldsymbol{f}$ of G belongs to $\boldsymbol{Z}(G)$ if, and only if, $S\boldsymbol{f} = S$ for every subgroup S of G, satisfying*

$$S \leq (G, n(x) < m(G)) \text{ or } F(0) \leq S,$$

where $m(G)$ is the invariant of condition (Ib) *and $F(0) = (2^\infty G, n(x) \leq 1)$ is the group of all those elements of G which are fixed elements of every proper automorphism of G.*

(IIb) $\boldsymbol{Z}(G)$ *is the direct product of $\boldsymbol{Z}^*(G)$ and the group $\boldsymbol{W}$ of order 2 which is generated by the automorphism $\boldsymbol{w}$ of G, satisfying:*

$x\boldsymbol{w} \left\{ \begin{matrix} = x \\ = x + v \end{matrix} \right\}$ *if* $x \left\{ \begin{matrix} \text{is} \\ \text{is not} \end{matrix} \right\}$ *an element of the join of $(G, n(x) < m(G))$ and $2^\infty G$, where v is the element $\neq 0$ in $F(0)$.*

Proof. 1. Assume that $\boldsymbol{f}$ belongs to $\boldsymbol{Z}(G)$.

If $G = D + G'$, then denote by $\boldsymbol{g}$ the automorphism of G, satisfying:

$$xg = x \text{ for } x \text{ in } D,$$
$$= -x \text{ for } x \text{ in } G'.$$

If u is any element of D, then $uf = d + u'$ with d in D and u' in G' and

$$ugf = uf = d + u'$$
$$= ufg = dg + u'g = d - u',$$

i. e. $2u' = 0$ and $2uf = 2d$ is an element of D and consequently:

(1) If f belongs to $\mathbf{Z}(G)$, D is a direct summand of G, then $2Df = 2D$.

2. Assume that f belongs to $\mathbf{Z}(G)$ and that $r(G, i) \neq 0$ for a certain integer i.

Then there exists by section 1 a decomposition

$$G = Z + G'$$

where Z is a cyclic group, generated by an element z of order p^t. If v is an element of G' such that $n(v) \leq i$, then $z + v$ generates a (cyclic) group Z' such that

$$G = Z' + G'.$$

Since, by (1), $2Zf = 2Z$, $2Z'f = 2Z'$, $2G'f = 2G'$, there exist integers c, c' such that

$$c'2z + c'2v = c'2(z + v) = 2(z + v)f = 2zf + 2vf = c2z + 2vf,$$

and since $2vf$ is an element of G', it follows that $c'2z = c2z$ and $2c \equiv 2c' \bmod p^t$ and therefore $2vf = 2cv$. Thus it has been proved:

(2) If f belongs to $\mathbf{Z}(G)$, then there exists for every integer i with $r(G, i) \neq 0$ an integer c_i such that

$$xf = c_i x \text{ for every } x \text{ in } 2(G, n(x) \leq i).$$

3. If f belongs to $\mathbf{Z}(G)$ and the orders of the elements of $G/p^\omega G$ are not bounded, then there exists an infinity of integers i such that $r(G, i) \neq 0$ and therefore every element x in $2G$ is mapped by f upon a multiple of x.

If the orders of the elements of $G/p^\omega G$ are bounded, then it follows from section 1, that

$$G = G' + p^\infty G,$$

where the orders of the elements in G' are bounded and G' is therefore a direct sum of cyclic groups. If p^m is the maximum order of the elements in G' and x an element in $p^\infty G$ such that $m < n(x)$, then x is contained in a direct summand S of type p^∞ of $p^\infty G$. Since $Sf = 2Sf = 2S = S$, there exists by (1) an integer $c(x)$ such that

$$xf = c(x)x \quad \text{and} \quad 2c(x) \equiv 2c_m \bmod p^m.$$

If g is any element of G, then $g = g' + g''$ with g' in G' and g'' in $p^\infty G$ and

$$\begin{aligned} 2g\boldsymbol{f} &= c_m 2g \quad \text{for} \quad n(g'') \leq n(g) \leq m, \\ &= c(g'')2g \quad \text{for} \quad m < n(g) = n(g'') \end{aligned}$$

and therefore every element x in $2G$ is mapped upon a multiple of x by $\boldsymbol{f}$. Thus it has been proved:

(3) If $\boldsymbol{f}$ belongs to $\mathbf{Z}(G)$, then $S\boldsymbol{f} = S$ for every subgroup S of $2G$.

Note that $2G = G$ for $p \neq 2$ and that therefore Theorem 5. 2 implies:

(3a) If $p \neq 2$, then the automorphism $\boldsymbol{f}$ of G belongs to $\mathbf{Z}(G)$ if, and only if, $S\boldsymbol{f} = S$ for every subgroup S of G.

4. Assume that $p = 2$ and that $\boldsymbol{f}$ belongs to $\mathbf{Z}(G)$. There exists by (3) and Theorem 5. 2, an integer 2-adic number $c = c(\boldsymbol{f})$ such that $2x\boldsymbol{f} = c2x$ for every element x in G.

Put $x\boldsymbol{h} = x\boldsymbol{f} - cx$ and this function $\boldsymbol{h}$ satisfies:

$n(x\boldsymbol{h}) \leq 1$ for every element x in G;
$(x + y)\boldsymbol{h} = x\boldsymbol{h} + y\boldsymbol{h}$, since $\boldsymbol{f}$ is an automorphism;
$x\boldsymbol{gh} = x\boldsymbol{gf} - cx\boldsymbol{g} = x\boldsymbol{fg} - cx\boldsymbol{g} = x\boldsymbol{hg}$ for every proper automorphism $\boldsymbol{g}$ of G, since $\boldsymbol{f}$ belongs to $\mathbf{Z}(G)$.

If $r(G, i) \neq 0$, then there exists by section 1 a decomposition:

$$G = Z + G'$$

where Z is a cyclic group, generated by the element z of order 2^i. If v is an element of G' such that $n(v) \leq i$, then a proper automorphism $\boldsymbol{g}$ of G is defined by

$$z\boldsymbol{g} = z + v, \qquad x\boldsymbol{g} = x \text{ for every } x \text{ in } G'.$$

Since $n(z\boldsymbol{h}) \leq 1$, $z\boldsymbol{h} = 2^s z + z'$ with $i - 1 \leq s \leq i$, z' in G' and therefore

$$\begin{aligned} z\boldsymbol{hg} &= 2^s z\boldsymbol{g} + z'\boldsymbol{g} = 2^s(z + v) + z' \\ &= z\boldsymbol{gh} = (z + v)\boldsymbol{h} = 2^s z + z' + v\boldsymbol{h}, \end{aligned}$$

and this proves:

(4) If $G = Z + G'$, where Z is a cyclic group of order 2^i, and $\boldsymbol{f}$ belongs to $\mathbf{Z}(G)$, then there exists an s such that $i - 1 \leq s \leq i$ and $x\boldsymbol{h} = 2^s x$ for every x in G' with $n(x) \leq i$.

If g is any element of G such that $n(g) < i$, then $g = 2kz + g'$ with g' in G' and $n(g') < i$, and consequently: $g\boldsymbol{h} = 0$ by (4). This proves:

(5) If $p = 2$, $r(G, i) \neq 0$, and $\boldsymbol{f}$ belongs to $\mathbf{Z}(G)$, then $x\boldsymbol{f} = cx$ for every x with $n(x) < i$.

If the orders of the elements of $G/2^\omega G$ are not bounded, then there exists

an infinity of integers i such that $r(G, i) \neq 0$ and therefore (5) and Theorem 5.2 imply:

(6) If $p = 2$ and the orders of the elements of $G/2^\omega G$ are not bounded, then the proper automorphism $\boldsymbol{f}$ of G belongs to $\boldsymbol{Z}(G)$ if, and only if,

$$S\boldsymbol{f} = S \text{ for every subgroup } S \text{ of } G.$$

5. Assume now that $p = 2$, that the orders of the elements of $G/2^\omega G$ are bounded and that $\boldsymbol{f}$ belongs to $\boldsymbol{Z}(G)$.

Then there exists by section 1 a decomposition:

$$G = G' + 2^\infty G$$

where the orders of the elements in G' are bounded and G' is therefore a direct sum of cyclic groups. If 2^m is the maximum order in G', then in using the notation, introduced in 4, it follows from (5) that

$x\boldsymbol{f} = cx$ for every element x, contained in the join of $2^\infty G$ and $(G, n(x) < m)$.

$$G' = Z + T + G'',$$

where $G'' \leq (G, n(x) < m)$, T either 0 or a direct sum of cyclic groups of order 2^m and Z a cyclic group, generated by an element z of order 2^m. Then

$$z\boldsymbol{h} = 2^s z + z' \text{ with } z' \text{ in } T + G'' + 2^\infty G \text{ and } m - 1 \leq s \leq m$$

and (4) implies:

$$x\boldsymbol{h} = 2^s x \text{ for every element } x \text{ in } T + G'' + 2^\infty G.$$

If $T \neq 0$, then reasons of symmetry imply that $z' = 0$ and therefore:

(7) If $r(G, m) \neq 1$, $2^\infty G = 0$, then

$$x\boldsymbol{f} = (c + 2^s)x \text{ for every element } x \text{ in } G.$$

If $2^\infty G \neq 0$, then there exists in $2^\infty G$ an element w of order 2^m and $2^s w = w\boldsymbol{h} = 0$ implies that $s = m$ and therefore that $x\boldsymbol{f} = cx$ for every x in G and thus it follows from (7) and Theorem 5.2:

(8) If $r(G, m) \neq 1$, then $\boldsymbol{f}$ belongs to $\boldsymbol{Z}(G)$ if, and only if, $S\boldsymbol{f} = S$ for every subgroup S of G.

If $r(G, m) = 1$, but $2^\infty G \neq 0$, then the above argument shows again that $s = m$, i. e. that

$$z\boldsymbol{h} = z' \text{ with } z' \text{ in } G'' + 2^\infty G.$$

If $\boldsymbol{g}$ is any automorphism of G, then

$$z\boldsymbol{g} = jz + z^* \text{ with } j \text{ an odd integer and } n(z^*) \leq m,\ z^* \text{ in } G'' + 2^\infty G.$$

Then

$$zgh = (jz + z^*)h = zh + z^*h = z', \text{ since } n(zh) \leq 1,$$
$$= zhg = z'g, \text{ i. e.}$$

zh is an element of $F(0)$. If conversely w is an element of $F(0)$ and the proper automorphism b is defined by

$$zb = cz + w, \qquad xb = cx \text{ for } x \text{ in } G'' + 2^\infty G,$$

where c is an integer 2-adic number $\not\equiv 0 \bmod 2$, then any proper automorphism g of G satisfies:

$$zbg = czg + wg = c(kz + z^*) + w \text{ with } k \text{ an odd number and } z^* \text{ in } (G'' + 2^\infty G, n(x) \leq m),$$
$$= k(cz + w) + cz^*, \text{ since } n(w) \leq 1 \text{ by Theorem 3. 1,}$$
$$= kzb + z^*b = zgb, \text{ and}$$
$$xbg = cxg = xgb \text{ for } x \text{ in } G'' + 2^\infty G, \text{ i. e.}$$

b belongs to $\mathbf{Z}(G)$. Thus it has been proved:

(9) If $r(G, m) = 1$, $2^\infty G \neq 0$, then the proper automorphism f of G belongs to $\mathbf{Z}(G)$ if, and only if, there exists an integral 2-adic number $c = c(f)$ such that

$$xf - c(f)x = 0 \text{ for every element } x \text{ in the join of } (G, n(x) < m) \text{ and } 2^\infty G,$$
$$\equiv 0 \bmod F(0) \text{ for every } x \text{ in } G.$$

If $r(G, m) = 1$, $2^\infty G = 0$, f an automorphism in $\mathbf{Z}(G)$, then, as proved above,

$$zh = 2^s z + z' \text{ with } z' \text{ in } G'', \text{ and } m - 1 \leq s \leq m.$$

If g is any automorphism of G, then $zg = jz + z^*$ with j odd and z^* in G'' and therefore

$$zgh = jzh + z^*h = 2^s z + z', \text{ since } n(z^*) < m,\ n(zh) \leq 1,$$
$$= zhg = 2^s zg + z'g = 2^s z + z'g, \text{ since } 2^s z \text{ is by Theorem 3. 1 an element of } F(0), \text{ i. e.}$$

$z' = z'g$ is an element of $F(0)$. It follows now from Theorem 3. 1 that $n(z') \leq 1$ and $z' = rz$ for a suitable r and thus it follows from Theorem 5. 2 that

(10) If $r(G, m) = 1$, $2^\infty G = 0$, then the proper automorphism f of G belongs to $\mathbf{Z}(G)$ if, and only if,

$$Sf = S \text{ for every subgroup } S \text{ of } G.$$

6. Theorem 3. 1, Theorem 5. 2, (9) and (10) imply:

(11) If $r(G, m) = 1$, $r(G, \infty) \neq 1$, then the proper automorphism $\mathfrak{f}$ of G belongs to $\mathbf{Z}(G)$ if, and only if,

$$S\mathfrak{f} = S \text{ for every subgroup } S \text{ of } G.$$

7. Assume now that $1 = r(G, m) = r(G, \infty)$ and $p = 2$. Then it follows from Theorem 3. 1 that

$$F(0) = (2^\infty G, n(x) \leq 1).$$

If $\mathfrak{f}$ belongs to $\mathbf{Z}(G)$ and either $F(0) \leq S$ or $S \leq (G, n(x) < m)$, then it follows from (9) that

$$S\mathfrak{f} = S.$$

If conversely the proper automorphism $\mathfrak{f}$ satisfies:

$S\mathfrak{f} = S$ for every subgroup S of G such that $F(0) \leq S$ or $S \leq (G, n(x) < m)$,

then there exists by Theorem 5. 2 an integer 2-adic number c such that

$x\mathfrak{f} = cx$ for every element x in $2G =$ join of $(G, n(x) < m)$ and $2^\infty G$.

If furthermore $S = Z + F(0)$, where Z is a cyclic direct summand of G, generated by an element z of order 2^m, then $S\mathfrak{f} = S$, i. e. $z\mathfrak{f} = jz + w$ where j is odd and w in $F(0)$. Since $2z\mathfrak{f} = c2z$, it follows that $2cz = 2jz$, i. e. $jz = (c + e2^{m-1})z$ with $e = 0, 1$. There exists furthermore in $2^\infty G$ an element u of order 2^{m+1}. Since w is an element of $F(0)$, since, by Theorem 3. 1, $F(0) = (2^\infty G, n(x) \leq 1)$, and since $r(G, \infty) = 1$, it follows that $w = d2^m u$ where $d = 0$ or 1. The cyclic group, generated by $z + u$, contains $F(0)$ and it follows from the assumption, concerning the automorphism $\mathfrak{f}$, that

$$\begin{aligned}(z + u)\mathfrak{f} &= k(z + u) = kz + ku \text{ where } k \text{ is an odd integer,}\\ &= z\mathfrak{f} + u\mathfrak{f} = jz + w + cu\\ &= (c + e2^{m-1})z + (c + d2^m)u.\end{aligned}$$

This implies that

$$\begin{aligned}k &\equiv c + e2^{m-1} \bmod 2^m\\ &\equiv c + d2^m \bmod 2^{m+1} \text{ and therefore}\\ &\equiv c \bmod 2^m.\end{aligned}$$

Consequently $e = 0$ and $cz = jz$ and now it follows from (9) that $\mathfrak{f}$ belongs to $\mathbf{Z}(G)$. This completes the proof of part (II) of the Theorem. Part (I) of the Theorem is a consequence of (3a), (6), (8) and (11).

Inspection of the automorphisms $\mathfrak{g}$, used in demonstrating the proposition (3) of the proof, shows that they are of order 2 and that therefore even the following statement holds true:

(12) If the automorphism f of the primary Abelian group G of characteristic p belongs to the kernel[17] $\boldsymbol{K}(G)$ of $\boldsymbol{P}(G)$, then $Sf = S$ for every subgroup S of $2G$. Thus, if $p \neq 2$, $\boldsymbol{K}(G) = \boldsymbol{Z}(G)$.

If $p = 2$, but the orders of the elements in G are not bounded, then there exist automorphisms in $\boldsymbol{Z}(G)$ of infinite order (cf. Theorem 5. 2 and the fact that there exist integral 2-adic numbers whose multiplicative order is infinite) and it follows from a general theorem on kernels[18] that $\boldsymbol{K}(G) = \boldsymbol{Z}(G)$.

COROLLARY.[19] *The group $\boldsymbol{P}(G)$ of all the proper automorphisms of the primary Abelian group G is Abelian if, and only if, either G is irreducible or G is a direct sum of a group of order 2 and a group of type 2^∞.*

Remark. There exist Abelian groups without elements of finite order whose group of automorphisms is Abelian, e. g. the additive group of the rational numbers and its subgroups, the additive group of the integral p-adic numbers. An example of a finite non-commutative group whose group of automorphisms is Abelian has been given by Miller.

Proof. If G is irreducible, then every automorphism of G maps every subgroup of G upon itself and belongs therefore to the centralizer of $\boldsymbol{P}(G)$. If $G = T + G'$ where T is of order 2 and G' of type 2^∞, then every automorphism f of G satisfies: $Sf = S$ for every subgroup S of G such that $F(0) \leq S$, since $G'f = G'$ and $Tf \leq T + F(0)$. It follows therefore from the Theorem that $\boldsymbol{Z}(G) = \boldsymbol{P}(G)$, if G satisfies the condition of the Corollary.

If $G = Z + Y + G'$ where Z and Y are irreducible groups neither of them of order 2, then it follows from the Theorem that there exists an automorphism f of G which does not belong to $\boldsymbol{Z}(G)$ and this completes the proof.[20]

THE INSTITUTE FOR ADVANCED STUDY,
PRINCETON, N. J.

---

[17] The kernel of the (non-commutative) group Q consists of all the elements x in Q, satisfying $x^{-1}Sx = S$ for every subgroup S of Q, i. e. $x^{-1}yx = yh^{(y,x)}$ for every element y in Q. Cf. the paper, mentioned in footnote [3].

[18] Cf. Satz 3 of the paper, mentioned in footnote [2].

[19] This proposition has been proved for finite Abelian groups by A. Chatelet, "Les groupes abéliens finis et les modules des points entiers," *Travaux et mémoirs de l'Université de Lille*, nouv. sér. II, 3. Paris, 1925.

[20] In applying similar methods it is not difficult to prove the following proposition: *If G is a primary Abelian group and $\boldsymbol{Z}(G) < \boldsymbol{P}(G)$, then $\boldsymbol{P}(G)/\boldsymbol{Z}(G)$ is Abelian if, and only if,*

$$G = U + V$$

where U is of order 2 and V a cyclic group of order 2^i with $1 < i$.

POINTWISE PERIODIC HOMEOMORPHISMS.[1]

By DEANE MONTGOMERY.

1. In considering periodic transformations of locally Euclidean spaces there is suggested the question of whether or not there may exist a transformation under which each point is periodic but which is such that the periods of the points have no finite upper bound. The purpose of this paper is to show that such transformations do not exist, that is to show that any transformation under which each point is periodic is itself periodic. The known results on transformations of period p therefore apply for some p to any transformation under which each point is periodic.[2]

2. The space M considered here is a connected metric space[3] each of whose points is included in an open set homeomorphic to the interior of a solid n-dimensional sphere. If T is any transformation of M into itself T^k denotes the k-th iterate of T. It is understood that $T^0 = I$. A point x of M is periodic if there exists some positive integer k such that $T^k(x) = x$. The smallest positive integer k having this property is called the period of x. If each point of M is periodic under T, then T is said to be pointwise periodic. If there is some positive integer k such that $T^k = I$, the identity, then T is said to be periodic.

THEOREM. *If T is a pointwise periodic homeomorphism of M into itself then T is periodic.*

The proof is based on a lemma. It is convenient to make use of the integer valued function $p(x)$ defined to be the period of x under T. This function is defined everywhere in M and is lower semi-continuous. Let K be the set of points where $p(x)$ has an infinite least upper bound. The set K is closed.

---

[1] Received May 1, 1936.

[2] The principal contributions to the theory of periodic transformations have been made by Brouwer, Kerékjártó, Eilenberg, Newman, and P. A. Smith.

The author wishes to say that he has benefited from conversations with Leo Zippin on the topic of this note.

[3] Metric spaces were first used by Fréchet. See his *Les Espaces Abstraits*. See also Kuratowski, *Topologie I*.

LEMMA. *On every connected open set R in $M - K$, $p(x)$ is bounded.*

Since $p(x)$ is lower semi-continuous and since M is locally complete, $p(x)$ must be continuous on a set C which is everywhere dense in M;[4] C is an open set because $p(x)$ is integer valued. In the vicinity of a point of C $p(x)$ must be constant. There therefore exists a connected open set H in R such that $p(x)$ is constant and equal to an integer k everywhere in H. If H is not all of R let b be a point of R on the boundary of H. There is a neighborhood $U(b)$ in R and an integer N such that for any point x in $U(b)$, $p(x)$ is at most N. This is due to the fact that $p(x)$ has a finite least upper bound at b. Now consider the transformation $S = T^k$. Let

$$H_1 = \sum_{i=0}^{N!} S^i[H + U(b)].$$

This set is transformed homeomorphically into itself by S. Furthermore S is periodic of period at most $N!$, and for every x in H, $S(x) = x$. Since H_1 is connected it can be concluded from Newman's results[5] that every point in H_1 is fixed under S. But this means that if x is any point in H_1, then $T^k(x) = x$, so that every point of H_1 has period at most k. The set H_1 may, by the same process, be enlarged to a set H_2 (also in R) every point of which has period at most k. Continuing in this way by transfinite induction of any order necessary there is finally obtained a set H_α which coincides with R and every point of which has period at most k. This completes the proof of the Lemma.

3. The proof of the theorem will now be given by means of a *reductio ad absurdum*. Assume that $p(x)$ is unbounded. It follows from this assumption and from the lemma that $M - K$ is not connected and therefore that K is not vacuous. Let a be a point of K where $p(x|K)$[6] is continuous. There is an open set H including a such that $p(x|K)$ is constant and equal to some integer k everywhere in $H \cdot K$. Let R be any component of $M - K$ having points in H. There is by the lemma an integer N such that every point of R has period at most N. Now consider the set

$$W = \sum_{i=0}^{N!} S^i(R)$$

---

[4] Kuratowski, *loc. cit.*, p. 189.

[5] *Quarterly Journal of Mathematics*, vol. 2 (1931), p. 1.

[6] This notation is used to indicate the function whose domain of definition is K and which is equal to $p(x)$ everywhere on this set.

where $S = T^k$. If any point of $\bar{R} \cdot K \cdot H$ is an inner point of $W + K$, $p(x)$ is bounded at that point. Let c be any point of $\bar{R} \cdot K \cdot H$; by the remark just made c is not an inner point of $W + K$. Let V be a connected open set which is in H and which contains c. The set $V - \bar{W}$ is not vacuous. The set $V - W$ is separated from W by K and each point of $K \cdot H$ is fixed under S, so that the transformation S^* now to be defined is a homeomorphism of $V + W$ into itself. If x is in W, $S^*(x) = S(x)$ and if x is in $V - W$, $S^*(x) = x$. The set $V + W$ is connected and locally Euclidean so that Newman's results may again be applied, and it may be concluded that every point of $V + W$ is fixed under S^*. Therefore every point of W and in particular every point of R has period at most k. It follows, since R is an arbitrary component of $M - K$ having points in H, that every point of $(M - K) \cdot H$ has period at most k. But the same is true for every point of $H \cdot K$. Hence at points of $H \cdot K$ $p(x)$ has a finite least upper bound, and this contradicts the definition of K. The theorem is now demonstrated.

SMITH COLLEGE.

TRANSLATION GROUPS OF THREE-SPACE.[1]

By DEANE MONTGOMERY and LEO ZIPPIN.

1. In this paper we seek conditions under which an abelian k-parameter group of transformations of an n-dimensional Euclidean space is essentially a k-parameter translation.[2] This problem is solved completely in case $k \geqq n-2$ or, in consequence, whenever $n=3$. For $n>3$ and general k, the problem is shown to be equivalent to an interesting question in topological product-spaces.

2. Let V^k denote the k-parameter (abelian) group of vectors of Euclidean k-dimensional space. The elements of V^k will be denoted by $u, v, w, \cdots$, the identity, in particular, by v_0. Let R denote an arbitrary metric space (which will subsequently be specialized). The points of R will be denoted by $x, y, z, \cdots$.

We shall say that V^k is a group of transformations of R provided that with each v there is uniquely associated a single-valued mapping of R into itself (*a priori*, not necessarily covering R) which we shall represent by

$$vx = x', \qquad x, x' \subset R;$$

furthermore these mappings must satisfy the following conditions:

i) $v_0 x = x$, for all x,

ii) $u(vx) = wx$, for all x, where $w = u + v$,

iii) vx is continuous in v and x simultaneously, that is if $v_n \to v$, and $x_n \to x$, then $v_n x_n \to vx$.[3]

If V denotes an arbitrary subset of V^k and X an arbitrary subset of R, then VX shall denote the set of all points in R of the form vx, where v is in V and x in X.

3. If x is any point of R, v any element of V^k, then for the element $y = (-v)x$ it follows from the first two conditions on V^k that $x = vy$. This implies that each v "covers" all of R and has, furthermore, a single-valued

[1] Received May 1, 1936.

[2] V. Niemytski has made such a study of one-parameter groups under the hypothesis of differentiability. See *Comptes Rendus de l'Académie des Sciences de Paris*, vol. 199 (1934), p. 18. Also, "Über Vollständige instabile dynamische Systeme," *Annali di Matematico* (4), vol. 14, pp. 275-286.

[3] Read "converges to."

inverse. The continuity condition assures that each V is a homeomorphism of R into itself.

Definition. A group of transformations V^k of R is called *dispersive* if with every pair of points x and y of R there is associated a pair of positive numbers ϵ and N such that for every point $z \subset S(x, \epsilon)$[4] and every v of norm greater than N, $| v | > N$, $vz \not\subset S(y, \epsilon)$.

4. THEOREM 1. *If V^k is a dispersive group of motions of R and R is locally compact, connected, and separable, then R may be represented as the topological product of two subspaces T_k and R_k:*

$$R = T_k \times R_k,$$

where T_k is homeomorphic to V^k.

We shall obtain the proof after a sequence of simple lemmas.

LEMMA 1. *If $x_n \to x$ and $v_n x_n \to y$, then $| v_n |$ is bounded.*

With each fixed $\epsilon > 0$ there is associated an integer $\bar{m}$ such that $x_n \subset S(x, \epsilon)$ and $v_n x_n \subset S(y, \epsilon)$, whenever $n \geqq \bar{m}$. For the ϵ and N which are associated with x and y by the dispersive character of V^k and for the corresponding $\bar{m}$, we have $| v_n | \leqq N$ whenever $n \geqq \bar{m}$.

LEMMA 2. *If $ux = vx$ for some point x, then $u = v$.*

Let $w = u - v$. Then $wx = (u - v)x = x$. Let $w_n = nw$, $n = 1, 2, \cdots$. Clearly $w_n x = x$. As a special case of Lemma 1, $| w_n |$ is bounded. Since $| w_n | = n | w |$, it follows that $| w | = 0$ and w must be the identity. That is, $u = v$.

LEMMA 3. *If $x_n \to x$ and $v_n x_n \to y$, then there is a v such that $v_n \to v$ and $y = vx$.*

Since v_n is bounded, by Lemma 1, there is necessarily at least one element v and a subsequence v_{i_n} such that $v_{i_n} \to v$. By the continuity, $v_{i_n} x \to vx$, so that $vx = y$. By Lemma 2, v is uniquely determined.

COROLLARY 1. *If V is any closed subset of V^k and x any point of R, Vx is closed.*

COROLLARY 2. *If $x_n \to x$ and $v_n x_n \to x$, then $v_n \to v_0$.*

LEMMA 4. *For fixed x, $V^k x$ is homeomorphic to the space V^k.*

The required correspondence between points of $V^k x$ and V^k is given by

[4] $S(x, \epsilon)$ denotes the set of points at a distance $\leqq \epsilon$ from x.

$$x' = v'x.$$

To each v', since x is fixed, there corresponds one and only one x'. To each x' of $V^k x$ there corresponds at least one v' such that $x' = v'x$ and, by Lemma 2, this v' is unique. The correspondence is therefore bi-univalued. That it is also bicontinuous is a consequence of Lemma 3 and the continuity iii).

4. Now let V denote a one-parameter closed subgroup of V^k. It follows, from definitions, that V is a dispersive group of transformations of R. Consequently all of the above Lemmas hold for the group V.

LEMMA 5. *Let v be any element of V, $v_t = tv$, $0 \leqq t \leqq c$, any constant. Then given any ϵ there exists a $\delta > 0$ such that if (x, y) [= distance from x to y] $< \delta$, then $(v_t x, v_t y) < \epsilon$.*

This is a trivial consequence of the continuity of ux in u and x simultaneously.

Since Vx is homeomorphic to a straight line, by Lemma 4, it is a consequence of Lemma 5 that the space R is decomposed into a regular family of curves in the sense of H. Whitney. We shall use his work on cross-sections.[5] We define a *true-section*: a set of points Z of R is said to be a *true-section* (with respect to the regular family of curves) provided that

1) for every y of R the set $Z \cdot Vy$ consists of at most a single point,

2) if z_n is any sequence of points of Z and v_n any sequence of elements, $n = 1, 2, \cdots$, such that for some y of R, $v_n z_n \to y$ then there must exist a point z of Z and a subsequence z_{i_n} such that $z_{i_n} \to z$.

The second of these conditions implies, at once, that Z is closed. For if a point $z \subset \bar{Z}$, there is a sequence of points z_n of Z such that $z_n \to z$. Then $v_0 z_n = z_n \to z$, and 2) assures us that there is a $z' \subset Z$ and a subsequence z_{i_n} such that $z_{i_n} \to z'$. Clearly, $z' = z \subset Z$. Condition 2) implies also that VZ is closed. For if y is a limit point of VZ then for an appropriate choice of points z_n of Z and elements v_n, $v_n z_n \to y$. Certainly $v_{i_n} z_{i_n} \to y$, for any subsequence (i_n), $n = 1, 2, \cdots$. Then if $z_{i_n} \to z$, for some z of Z, we know from Lemma 3 that $y = vz$ for some v. It should be remarked that 2) is stronger than the requirement that Z and VZ be closed.[6] On the other hand, if Z is *compact* and closed it will obviously satisfy 2).

---

[5] H. Whitney, *Annals of Mathematics*, vol. 34 (1933), pp. 244-270. As Whitney points out in the last paragraph of his paper, cross sections may be obtained quite easily when, as here, the curves are given by a one-parameter group.

[6] The distinction can be made clear in a simple example. Let E be the Euclidean plane, V the group of translations parallel to the x-axis and Z the set of points consisting of the origin of coördinates together with the points of the hyperbola $xy = 1$. Here Z is *not* a true-section in the sense of condition 2.

A true section Z is said to be a true section at a point x provided that for all points y sufficiently near to x the intersection $Z \cdot Vy$ is not vacuous. It is said to "cover" a point y provided VZ covers y in the usual sense. It is clear that a true section at a point x "covers" x and also all points sufficiently close to x. Two true-sections Z_1 and Z_2 will be said to overlap provided the intersection $(VZ_1) \cdot (VZ_2)$ is not vacuous. A true-section Z "contains" a true-section Z' if VZ contains Z'. In case Z contains Z' in the customary sense, Z will be said to extend Z'.

5. LEMMA 6. *True sections exist at every point x of R.*

Let W denote a Whitney cross-section at the point x.[7] In virtue of Corollary 2 to Lemma 3, there must exist a positive ϵ such that for all $y \subset S(x, \epsilon)$, W satisfies the first condition for a true-section. By reason of being a Whitney section W, for some ϵ' which may be supposed less than ϵ has the property that for all $y \subset S(x, \epsilon')$, $W \cdot Vy$ is not vacuous. Since R is supposed *locally compact*,[8] for some $\epsilon'' < \epsilon'$, $S(x, \epsilon'')$ is compact. Now, $Z = W \cdot S(x, \epsilon'')$ is a true-section at the point x.

LEMMA 7. *Given two overlapping true-sections Z and Z', there exists a true section Z'' which "contains" Z and is an "extension" of Z'.*[9]

Since Z and VZ' are closed (see no. 4) it follows that $X = Z \cdot VZ'$ is closed. With each point x of X there is associated a unique element v and point x' of Z' such that $vx = x'$. The element v is a continuous function of x. For if x_n is a sequence of points of X such that $x_n \to x$, let $x'_n = v_n x_n$ be the corresponding points of Z'. Since $(-v_n)x'_n \to x$, it follows from condition 2) of no. 4 that a subsequence of the x'_n must converge to a point x', and this point must belong to Z' which is closed. By Lemma 3, $x' = vx$, where $v_n \to v$. Now v as a continuous *real-valued* function [10] of a closed subset X of a metric space Z can be extended to a continuous function over all of Z.[11] The set of points vz, where z ranges over Z and v is the associated functional value is easily seen to be the desired true section Z''.

6. We may proceed, finally, to a proof of Theorem 1. With each point x of R we may associate a true section Z_x. Each Z_x "covers" the corresponding point x. Since R is assumed separable, there will exist a countable sequence $Z_1, Z_2, \cdots$, of these sets such that they "cover" space: i. e. such that VZ_n,

[7] *Loc. cit.*, p. 256.

[8] This is the only point at which the local compactness of R is used.

[9] See definitions in no. 4.

[10] Since V is a one-parameter vector group.

[11] See for example Alexandroff and Hopf, *Topologie* I, p. 73.

$n = 1, 2, \cdots$, together cover R. Because R is connected we may assume that these sections are ordered in such a way that VZ_{n+1} has a non vacuous intersection with $\sum_{i=1}^{n} VZ_i$. By Lemma 7 there exists a true-section Z^1 which "extends" Z_1, and "contains" Z_2. By the same lemma there exists a true-section Z^2 which "extends" Z^1 and contains Z_2. Proceeding inductively we construct a set R_1 which is the point set sum of the monotonic increasing sequence of true-sections $Z^1, Z^2, Z^3, \cdots$. It is immediate that R_1 satisfies condition 1) of a true-section. Now suppose that y is any point such that for a sequence of points z_n of R_1 and a proper choice of elements v_n, $v_n z_n \to y$. Let Z_k be a true section in our first sequence which covers y. Then, for sufficiently large n, it will also cover all z_n. Now Z_k is "contained" in Z^k and Z^k must contain almost all z_n. But since Z^k is a true-section condition 2) must be satisfied for it and therefore condition 2) is also satisfied for R_1. It follows that R_1 is a true section.

It is trivial that R may be expressed as the topological product of V and R_1: $R = V \times R_1$ since each point of R is in the form vz for some unique element v and unique point z of R_1. Since R_1 is a closed subset of R, it is itself a locally compact separable metric space.

Now if $k = 1$, in our original theorem, the proof is complete. If $k > 1$, let V^{k-1} be a vector subgroup of V^k orthogonal to the original V of the argument above. We shall show that it is possible to interpret V^{k-1} as a dispersive group of motions over R_1. To this end, let v be an arbitrary element of V^{k-1}, x be any point of R_1 and $y = vx$ the corresponding point of R. The point y belongs to one and only one Vz where z is a point of R_1. We define a new correspondence (essentially a "projection" of the original one)

$$v: \quad \tilde{v}x = z.$$

Here the point z is of the form ty for a unique element t of the group V.

That these new correspondences satisfy the conditions i) and iii) for a group of motions over R_1 (see no. 2) is clear. That they satisfy ii) follows from the obvious fact that $\tilde{u}(\tilde{v}x) \subset V\tilde{w}x$, where the associated $w = u + v$. Since R_1 is a true section this implies that $\tilde{u}(\tilde{v}x) = \tilde{w}x$, both being points of R_1. We have left to verify that V^{k-1}, thus interpreted, is dispersive. Let x and y denote arbitrary points of R_1 and let ϵ and N be chosen in accordance with the definition in no. 3. Now if there is any point z in $S(x, \epsilon)$ and any v of V^{k-1} such that the $\tilde{v}z \subset S(y, \epsilon)$ then, for the appropriate t of V, $(t + v)z \subset S(y, \epsilon)$. But then, $| t + v | \leqq N$ and therefore $| v | \leqq N$ since v and t are orthogonal.

If we now choose in V^{k-1} a closed one-parameter vector group V', it follows by what has already been shown that

$$R_1 = V' \times R_2,$$

for some appropriately constructed "true-section" R_2 of R_1. By a finite induction, and because of the associativity of topological products, $R = V^k \times R_k$, for the corresponding R_k.

Since V^k is k-dimensional and homeomorphic to a subset of R it follows that $k \leqq \dim R$. In case $\dim R = k$, it follows from a theorem due to Hurewicz [12] that $\dim R_k = 0$.

7. *Definition.* We shall say that a V^k is a k-parameter translation group of a Euclidean space E_n if V^k is a group of motions of E_n and if there is a coördinate system $x_1, \cdots, x_n$ in E_n and a generating set of k vectors $v_1, v_2, \cdots, v_k$ in V^k such that to the vector rv_i, where r is an arbitrary real number, there corresponds the translation:

$$x'_j = x_j + \delta_j{}^i r x_j, \qquad (j = 1, 2, \cdots, n).$$

THEOREM 2. *The following three propositions are equivalent, n being arbitrary but fixed:*

A) A dispersive group of motions V^k of an E_n is a k-parameter translation group of E_n for all k;

B) If $E_n = E_k \times R_k$, then R_k is homeomorphic to E_{n-k}, for all k;

C) If $E_n = E_1 \times R_1$, then R_1 is homeomorphic to E_{n-1}.

Proof. That C) implies B) is a consequence of the associativity of topological products. That B) implies A) is a consequence of our Theorem 1. For from this theorem it follows that $E_n = V^k \times R_k$ and, since V^k is homeomorphic to E_k, B) implies that R_k is homeomorphic to E_{n-k}. Since this is the case we can introduce a coördinate system into E_n having the desired properties.

Finally, A) implies C). Suppose we have $E_n = E_1 \times R_1$. We may regard E_1 as a one-parameter vector group, some point e_0 corresponding to the identity vector and some point e_1 to the generating vector. Let z_0 be an arbitrary fixed point in R_1. Then E_1 and R_1 may be interpreted as subspaces of E_n, R_1 being the set of all points (e_0, z) where z ranges over R_1 and E_1 being the set of all points (e, z_0) where e ranges over E_1. Now corresponding to the vector e of E_1 we have in E_1 a certain translation; $e\colon e'(e) = e' + e$. We can extend this to a motion of all of E_n, the point (e', z) going over into the point $[(e' + e), z]$. It is quite easy to verify that E_1 now becomes a dispersive group of motions of E_n. Now, applying A) we get a coördinate

[12] W. Hurewicz, *Annals of Mathematics*, vol. 36 (1935), p. 194.

system in E_n such that the vectors of E_1 correspond to a change in the first coördinate only. Furthermore any desired change in the first coördinate can be accomplished by a suitable vector e of E_1. From this fact we may conclude that between the points of R_1 and the coördinates $(x_2, x_3, \cdots, x_n)$ there is a bi-univalued and bicontinuous correspondence.

Given any point (e_0, z) of R_1 we associate with it the set of its last $n-1$ coördinates: $x_2, \cdots, x_n$. No other point (e_0, z') of R_1 can have the same last coördinates for otherwise there would exist an element e of E_1 such that $[(e_0 + e), z']$ coincided with (e_0, z) (by suitably changing the first coördinate). But this contradicts the fact that $[(e_0 + e), z']$ cannot be a point of R_1 unless e is the identity e_0. On the other hand to any set of coördinates $(x_2, \cdots, x_n)$ there corresponds the unique point $(0, x_2, \cdots, x_n)$ and to this point (e, z) of E_n a unique point (e_0, z) of R_1. That the given correspondence is bi-univalued and bicontinuous should now be obvious.

Regarding C) we may remark that it is easy to establish a number of topological properties of R_1 which tend to corroborate the conjecture that it is indeed an E_{n-1}. Only in the case that $n \leqq 3$ are we able to prove this. We shall call attention here to the fact, which we shall have occasion to use in a moment, that if $E_n = E_k \times R_k$, then R_k is a locally compact separable metric space, connected and locally connected.

8. THEOREM 3. *A necessary and sufficient condition that a V^k $(k \geqq n-2)$ be a k-parameter translation group of E_n is that it be a dispersive group of motions of E_n.*

The necessity is quite obvious; we prove the sufficiency. Let x be any point of E_n and $V^k x$ the corresponding closed subset determined by the dispersive group V^k. This set is homeomorphic to an E_k. It is clear that if we adjoin to E_n the point at infinity, it becomes an n-sphere E^*_n and $V^k x$ becomes a k-sphere E^*_k. Then if $k = n-1$, E^*_k separates E^*_n into two connected sets and if $k = n-2$, E^*_k does not separate E^*_n (by the Alexander Duality Theorem). We can conclude at once that if $k = n-1$, $V^k x$ separates E_n and if $k = n-2$ it does not. Now we know that E_n can be expressed as $V_k \times R_k$. From what we have said it follows that for $k = n-1$, each point of R_k separates R_k into two connected sets and from this it follows that R_k is homeomorphic to an E_1.[13] In this case, compare no. 7, our theorem is proved.

Now if $k = n-2$, we see that no point of R_k is a cut-point and therefore R_k contains simple closed curves.[14] Let J be any simple closed curve of R_k.

[13] See, e. g., R. L. Moore, "Foundations of point-set theory," *American Mathematical Society Colloquium Publications*, vol. 13 (1932), p. 103 ff.

[14] R. L. Moore, *loc. cit.*, p. 124.

The set V^kJ is homeomorphic to the product of a circle by an E_k and it is easily seen that regarded as a subset of E^*_n it is the homeomorph of an absolute $(n-1)$-cycle. The cycle in question is obtained from the topologic product of a circle and an $(n-2)$-cube in which the point set corresponding to the product of the circle by the $(n-3)$-sphere bounding the cube has been "identified" with a single point. We conclude, by the Duality Theorem, that the set V^kJ separates E_n. Let x' and y' be two points in different components of $E_n - V^kJ$. If the points x and y of R_k belonging to V^kx' and V^ky' respectively belonged to the same component C of $R_k - J$, it would follow that x' and y' belonged to the connected set V^kC in $E_n - V^kJ$. The contradiction shows that J separates R_k. That no arc L of J separates R_k is a consequence of the Duality Theorem and the fact that V^kL contains no absolute $(n-1)$-cycle. We may conclude that R_k is a "cylinder-tree":[15] i. e. that it is homeomorphic to the point set $S_2 - B$ where S_2 is the 2-sphere and B some closed and totally disconnected subset of S_2. Since R_k is obviously not compact, it is clear that B is not vacuous. All that remains to show is that B cannot contain more than a single point. Now if B contains as many as two points, it is clear that $S_2 - B$ contains some simple closed curve J' such that J' does not bound any singular 2-cell on $S_2 - B$. In this case the corresponding simple closed curve J of R_k cannot bound any singular 2-cell of R_k. But it is clear that if J is any simple closed curve of R_k it must certainly bound a singular 2-cell C' in E_n. Now each point of V^kC' determines a unique point of R_k. The set of these points is obviously a continuous image of C' and therefore itself a singular 2-cell which is a subset of R_k. The contradiction completes the demonstration that R_k is homeomorphic to an E_2. In view of no. 7, our theorem is proved for $k = n-2$.

The case $k = n$ is instantly disposed of. In this case $E_n = V^n \times R_n$. If x denotes any point of R_n, V^nx being n-dimensional must contain inner points of E_n[16] and it follows at once that V^nx is open. Since it is also closed it must coincide with E_n. The set R_n is the single point x which obviously corresponds to the origin of the n-parameter translation group.

Finally we may state the obvious corollary to this theorem that for $n = 3$ any V^k (k necessarily $\leqq n$) of motions of E_n is a k-parameter translation group of E_n if and only if it is dispersive.

SMITH COLLEGE,
INSTITUTE FOR ADVANCED STUDY.

---

[15] L. Zippin, *American Journal of Mathematics*, vol. 52 (1930), pp. 331-350.

[16] K. Menger, *Dimensionstheorie*, Teubner, 1928.

DIFFERENTIAL PROPERTIES OF ABSTRACT TRANSFORMATION GROUPS WITH ABSTRACT PARAMETERS.[1]

By A. D. MICHAL and V. ELCONIN.

Introduction. Our main object in this paper is to present first-order differential equations which are satisfied by and, under specified conditions, characterize certain generalizations of abstract differentiable transformation groups. Our results clarify the classical theory of Lie, and abstract his first fundamental theorem and several related theorems to transformations in Banach spaces. A 'structural function,' analogous to the classical structural constants, appears in the condition of complete integrability for the fundamental differential equation and is shown to relate this equation to the differential equations of the so-called parameter groups in the same way that the structural constants relate the corresponding classical equations. Jacobi's identity is abstracted to Banach spaces and is used to establish an important property of the structural function. A differentiable group of matrix transformations and a group of linear functional transformations are used to illustrate the preceding results. Finally, several important unsolved problems are indicated, which arise when the differential equations are given and we try to determine when, in what domains, and with what group properties solutions of the equations exist.

The main definitions are collected in section 1. In section 2 the fundamental differential equation and the associated differential equations of the parameter groups are obtained, and conditions are given under which the solutions of these equations necessarily have group properties. The structural function is studied in section 3. The illustrative groups of transformations are in section 4.

1. *Definitions.* Let Δ, $\mathfrak{D}$ be any non-empty sets [2] and $=$ any reflexive, symmetric, transitive relation, such as logical identity, in Δ and in $\mathfrak{D}$.

DEF. 1.1. A mark T is a *transformation of* $\mathfrak{D}$ if Tx is in $\mathfrak{D}$ for all x in $\mathfrak{D}$.

DEF. 1.2. If T is a transformation of $\mathfrak{D}$, then T is *transitive over* $\mathfrak{D}$

[1] Presented to the Society, November 30, 1935 and April 11, 1936; received by the Editors, June 12, 1936.

[2] The sets Δ, $\mathfrak{D}$ will be in Banach spaces after Definition 1.8.

if for any x in $\mathfrak{D}$, $x = Ty$ for some y in $\mathfrak{D}$; T is *simple over* $\mathfrak{D}$ if $Tx = Ty$ implies $x = y$ for any x, y in $\mathfrak{D}$; T is *simply transitive over* $\mathfrak{D}$ if it is simple and transitive over $\mathfrak{D}$.

DEF. 1.3. If T, U are transformations of $\mathfrak{D}$, then $T = U$ *over* $\mathfrak{D}$ if

$$Tx = Ux \tag{1.1}$$

for all x in $\mathfrak{D}$; and TU, *the product of* T *and* U, is a transformation of $\mathfrak{D}$ satisfying

$$TUx = T(Ux) \tag{1.2}$$

for all x in $\mathfrak{D}$.

Transformation multiplication is evidently associative.

Let μ be in Δ. Let α' and $\alpha\beta$ be in Δ and T_α a transformation of $\mathfrak{D}$ for all α, β in Δ.

DEF. 1.4. If

$$T_\beta T_\alpha = T_{\alpha\beta} \text{ over } \mathfrak{D} \tag{1.3}$$

for all α, β in Δ, then T_α is *closed over* $\Delta, \mathfrak{D}$ *with respect to* $\alpha\beta$, and $\alpha\beta$ is a *parameter function over* $\Delta, \mathfrak{D}$ *of* T_α.

DEF. 1.5. If (1.3) holds and

$$\alpha(\beta\gamma) = (\alpha\beta)\gamma \tag{1.4}$$

for all α, β, γ in Δ, then T_α *generates a semi-group over* $\Delta, \mathfrak{D}$ *with respect to* $\alpha\beta$.

DEF. 1.6. If

$$T_\mu x = x \tag{1.5}$$

for all x in $\mathfrak{D}$, then μ is *a unit in* Δ *over* $\mathfrak{D}$ *of* T_α.

DEF. 1.7. If (1.3), (1.4), (1.5) hold, then T_α *generates an integral semi-group over* $\Delta, \mathfrak{D}$ *with respect to* $\alpha\beta$, μ.

DEF. 1.8. If (1.3), (1.4), (1.5) hold and

$$\alpha\mu = \alpha;\ \alpha\alpha' = \mu;\ \alpha\beta = \alpha\gamma \text{ implies } \beta = \gamma, \tag{1.6}$$

for all α, β, γ in Δ, then T_α *generates a group over* $\Delta, \mathfrak{D}$ *with respect to* $\alpha\beta, \mu, \alpha'$.

(1.4) and (1.6) implies that Δ is a group, in the usual sense, with $\alpha\beta$ as the group operation, μ as unit element, and α' as the inverse of α.

Let T be the set of transformations T_α for α in Δ. If the conditions in

Definitions 1. 5, 1. 7, or 1. 8 are satisfied, then T is respectively the semi-group, the integral semi-group, or the group which T_α generates, where in each case the product of elements in T is that defined in (1. 2), T_μ is the unit element of T in the second and third cases, and $T_{\alpha'}$ is the inverse of T_α in the third case. Since we are more concerned here with analytic properties of $T_\alpha x$ than with set-theoretical properties of T, we keep T_α rather than T in evidence.

The theory of T is equivalent to a theory of a basic function $T(\alpha, x)$ on $\Delta, \mathcal{D}$ to $\mathcal{D}$, since there is a 1 — 1 correspondence between the elements of T and the functions $T_\alpha x$ of x. However, we have preferred the transformation symbolism for several reasons, of which the most practical is the compact representation of iterations of the basic function furnished by transformation products.

Note that in the above definitions α, β can be replaced by any two distinct marks not already in use without altering the meaning of the defined propositions. Marks with this property will be called *umbrals.*

If the range and domains of $F(x, y, \cdots, z)$ are in Banach spaces [3] $d^{xy\cdots z}_{\xi\eta\cdots\zeta} F(x, y, \cdots, z)$ will denote the Fréchet differential [4] of $F(x, y, \cdots, z)$ with increments $\xi, \eta, \cdots, \zeta$. This form is occasionally ambiguous and we use $d^{xy\cdots z}_{\rho\sigma\cdots\tau} F(\rho, \sigma, \cdots, \tau)$ where $\rho, \sigma, \cdots, \tau$ are umbrals. $d^x_\xi F(x, y, \cdots, z)$ is sometimes more conveniently written as $F(x, y, \cdots, z; \xi)$. Successive differentiation with respect to the first argument x may accordingly be expressed by the adjunction of a succession of semi-colons and increments.

Now suppose $\Delta, \mathcal{D}$ are domains (open sets) in Banach spaces Σ, S.

DEF. 1. 9. T_α is *continuous in* α *over* $\Delta, \mathcal{D}$ or *continuous over* $\Delta, \mathcal{D}$ according as $T_\alpha x$ is continuous in α or in α, x for all α, x in $\Delta, \mathcal{D}$.

T_α is *differentiable in* α over $\Delta, \mathcal{D}$ or *differentiable over* $\Delta, \mathcal{D}$ *according* as $d^\alpha_\xi T_\alpha x$ exists or $d^{\alpha x}_{\xi z} T_\alpha x$ exists for all α, x in $\Delta, \mathcal{D}$.

DEF. 1. 10. If T_α is differentiable in α over $\Delta, \mathcal{D}$, then α is *essential over* $\Delta, \mathcal{D}$ *in* T_α if for any function $\eta(\alpha)$ on Δ to Σ

$$d^\alpha_{\eta(\alpha)} T_\alpha x = 0$$

for all α, x in $\Delta, \mathcal{D}$ implies $\eta(\alpha) = 0$ for all α in Δ.

DEF. 1. 11. A function F whose range and domains are in Banach

[3] S. Banach, *Théorie des Opérations Linéaires* (1932).

[4] M. Fréchet, *Annales Sc. Ec. Norm. Sup.*, t. 42 (1925), pp. 293-323. See also T. H. Hildebrandt and L. M. Graves, *Transactions of the American Mathematical Society*, vol. 29 (1927); A. D. Michal, *Annali di Matematica* (in press).

spaces is *definite* in one of its arguments with respect to the others if $F = 0$ for all values of the latter implies that the former $= 0$.

In Definitions 1. 9, 1. 10, α is an umbral.

2. *The fundamental differential equation.* Let $\Delta, \mathfrak{D}$ be domains in Banach spaces Σ, S. Let $\alpha\beta, \alpha'$ be in Δ and let T_α be a transformation of $\mathfrak{D}$ for all α, β in Δ. For brevity below, let α, β, γ be in Δ; ξ, η in Σ; x in $\mathfrak{D}$; z in S.

THEOREM 2. 1. *If*

(i) T_α *is closed over* $\Delta, \mathfrak{D}$ *with respect to* $\alpha\beta$,

(ii) T_α *is differentiable over* $\Delta, \mathfrak{D}$ *and* $d_z^x T_\alpha x$ *is solvable in* z *for all* α, x,

(iii) $\dfrac{\beta}{\alpha}$ *exists in* Δ, *differentiable in* α, *and satisfying*

$$\alpha \frac{\beta}{\alpha} = \frac{\alpha\beta}{\alpha} = \beta$$

for all α, β, *then for all* α, x, ξ

$$d_\xi^\alpha T_\alpha x = U\big(T_\alpha x, V(T_\alpha x, \Omega(\alpha, \xi))\big), \tag{2.1}$$

where $U(x, z)$, $V(x, \xi)$, $\Omega(\alpha, \xi)$ *are on* $\mathfrak{D}, S$ *to* S; $\mathfrak{D}, \Sigma$ *to* S; *and on* Δ, Σ *to* Σ *respectively, are linear in their second arguments, and* $U(x, z)$ *is solvable in* z.

Proof. Let

$$d_{\xi z}^{\alpha x} T_\alpha x = G(\alpha, x, \xi) + H(\alpha, x, z).$$

Then

$$d_\xi^\alpha T_{\beta/\alpha} T_\alpha x = G\left(\frac{\beta}{\alpha}, T_\alpha x, d_\xi^\alpha \frac{\beta}{\alpha}\right) + H\left(\frac{\beta}{\alpha}, T_\alpha x, d_\xi^\alpha T_\alpha x\right) = 0,$$

since $T_{\beta/\alpha} T_\alpha = T_\beta$ over $\mathfrak{D}$. Let $P(\alpha, x, z)$ be the inverse with respect to z of $-H(\alpha, x, z)$. Then, with δ in Δ,

$$d_\xi^\alpha T_\alpha x = P\left(\frac{\beta}{\alpha}, T_\alpha x, G\left(\frac{\beta}{\alpha}, T_\alpha x, d_\xi^\alpha \frac{\beta}{\alpha}\right)\right) = P\left(\delta, T_\alpha x, G\left(\delta, T_\alpha x, d_\sigma^\alpha \frac{\alpha\delta}{\sigma}\right)\right),$$

where β has been replaced by $\alpha\delta$, since the middle member is obviously constant in β. Equation (2. 1) now follows, with

$$U(x, z) = P(\delta, x, z), \quad V(x, \xi) = G(\delta, x, \xi), \quad \text{and} \quad \Omega(\alpha, \xi) = d_\sigma^\alpha \frac{\alpha\delta}{\sigma}.$$

The solvability in z of $U(x, z)$ is evident from that of $H(\alpha, x, z)$.

COROLLARY 2. 1. *If in addition to the hypotheses of Theorem* 2. 1, T_α *is transitive over* $\mathfrak{D}$ *and* $d_\xi^\alpha T_\alpha x$ *is solvable in* ξ *for all* α, x, *then* $V(x, \xi)$, $\Omega(\alpha, \xi)$, *and* $V(x, \Omega(\alpha, \xi))$ *are solvable in* ξ.

Proof. The solvability in ξ of $V(x, \xi)$ is evident. That of

$$U\big(x, V(x, \Omega(\alpha, \xi))\big)$$

follows from (2. 1) and the transitivity of T_α. The remaining conclusions are consequences of Lemma 2. 1 below, which can be stated briefly with the aid of

DEF. 2. 1. Let E, F, G be Banach spaces. A mark L is *a linear transformation of E to F* if Lx is a linear function on E to F. If L, K are linear transformations of E to F and of F to G, then KL is the linear transformation of E to G satisfying

$$KLx = K(Lx)$$

for all x in E; if Lx is solvable and $L'x$ is its inverse, then L is *solvable* and the linear transformation L' on F to E is the *inverse of* L.

LEMMA 2. 1. *If k, n are positive integers, $k \leq n$; E_i is a Banach space for $i = 1, 2, \cdots, n+1$; for $i = 1, 2, \cdots, n$, L_i is a linear transformation of E_i to E_{i+1} and L'_i, the inverse of L_i, exists if $i \neq k$; then L_k is solvable if and only if $L = L_n L_{n-1} \cdots L_1$ is solvable.*

Proof. If L_k is solvable, so is L, since the product of solvable linear transformations is solvable. If L is solvable, so is L_k, since

$$L'_k = L_{k-1} L_{k-2} \cdots L_1 L' L_n L_{n-1} \cdots L_{k+1}.$$

COROLLARY 2. 2. *If* (i) *and* (ii) *of Theorem* 2. 1 *are satisfied, and for all α, β*

$$\alpha(\alpha'\beta) = \alpha'(\alpha\beta) = \beta$$

and $\alpha'\beta$ is differentiable in α, then the conclusion of Theorem 2. 1 *holds.*

Proof. In the proof of Theorem 2. 1, replace $\frac{\beta}{\alpha}$ by $\alpha'\beta$.

Let $\alpha\beta = \Lambda_\beta \alpha = \Pi_\alpha \beta$. Then Λ_β and Π_α are transformations of Δ.

THEOREM 2. 2. *If*

(i) *Δ is a group with $\alpha\beta$ as group operation, μ as unit element, and α' as the inverse of α for all α,*

(ii) *$\alpha\beta$ and α' are differentiable in α for all α, β, then $\alpha\beta$ is differentiable in β for all α, β; Λ_β and Π_α generate groups over Δ, Δ with respect to $\alpha\beta, \mu, \alpha'$; β and α are essential over Δ, Δ in Λ_β and Π_α respectively; Λ_β and Π_α are simply transitive over Δ for all α, β.*

Proof. Since $\alpha\beta = (\beta'\alpha')'$ and $\alpha\beta, \alpha'$ are differentiable in α, $\alpha\beta$ is dif-

ferentiable in β. Λ_β and Π_α are closed over Δ, Δ with respect to $\alpha\beta$; hence the second conclusion follows from the group properties of Δ, as does the fourth conclusion. The third conclusion follows from the solvability and differentiability of $\alpha\beta$ in α and in β, and from

LEMMA 2.2. *If a, b are in Banach spaces E, F; the functions $G(x)$, $H(y)$ are on a neighborhood of a to F and on a neighborhood of b to E; $d^a_\mu G(a)$ and $d^b_\nu H(b)$ exist; and*

$$H(G(x)) = x, \qquad G(H(y)) = y \tag{2.2}$$

for all x, y in neighborhoods of a, b; then the linear functions $d^a_\mu G(a)$, $d^b_\nu H(b)$ of μ, ν are mutually inverse.

Proof. Let $L\mu = d^a_\mu G(a)$, $K\nu = d^b_\nu H(b)$. Then differentiating equations (2.2) at $x = a$, $y = b$ gives

$$KL\mu = \mu, \qquad LK\nu = \nu$$

for all μ, ν in E, F.

Let T, Λ, Π be the sets of the transformations T_α, Λ_β, Π_α.

DEF. 2.2. If $\alpha\beta$ is a parameter function over $\Delta, \mathcal{D}$ of T_α, then Λ_β, Π_α are respectively the *first* and *second transformations* of Δ *associated with* T_α. If premise (i) of Theorem 2.2 holds, then Λ, Π are respectively the *first* and *second parameter groups associated with* T.

LEMMA 2.3. *If T_α is differentiable in α over $\Delta, \mathcal{D}$, α is essential over $\Delta, \mathcal{D}$ in T_α, the conclusion of Theorem 2.1 holds, and $\Omega(\alpha, \xi)$ is solvable in ξ for all α, then $U(x, V(x, \xi))$ is definite in ξ with respect to x.*

Proof. Suppose the conclusion is false. Then for some $\zeta \neq 0$ in Δ, $U(x, V(x, \zeta)) = 0$ for all x. Let $\eta(\alpha) = \Omega'(\alpha, \zeta)$, where $\Omega'(\alpha, \xi)$ is the inverse with respect to ξ of $\Omega(\alpha, \xi)$. Then $\eta(\alpha) \neq 0$, and from (2.1)

$$d^\alpha_{\eta(\alpha)} T_\alpha x = U\big(T_\alpha x, V(T_\alpha x, \Omega(\alpha, \eta(\alpha)))\big) = U(T_\alpha x, V(T_\alpha x, \zeta)) = 0$$

for all α, x, contrary to the essentiality of α in $T_\alpha x$.

THEOREM 2.3. *If*

(i) *T_α is closed over $\Delta, \mathcal{D}$ with respect to $\alpha\beta$, differentiable over $\Delta, \mathcal{D}$, transitive over $\mathcal{D}$ for all α, α is essential over $\Delta, \mathcal{D}$ in T_α, and $d^x_z T_\alpha x$ is solvable in z for all α, x,*

(ii) *α' is differentiable in α and such that for all α, β, ξ*

$$\alpha(\alpha'\beta) = \alpha'(\alpha\beta) = \beta, \qquad (\alpha\beta)' = \beta'\alpha', \qquad (\alpha')' = \alpha,$$

and $d^{\alpha}_{\xi}\alpha'\beta$ *exists and is solvable in* ξ, *then the conclusion of Theorem* 2.1 *holds, and for all* α, β, ξ, η, $\Omega(\alpha, \xi)$ *is solvable in* ξ, $\alpha\beta$ *is differentiable in* α *and in* β, *and*

$$d^{\beta}_{\eta}\alpha\beta = \Omega'(\alpha\beta, \Omega(\beta, \eta)), \tag{2.3}$$

$$d^{\alpha}_{\xi}\alpha\beta = \omega\big((\alpha\beta)', \Omega'\big((\alpha\beta)', \Omega(\alpha', \omega(\alpha, \xi))\big)\big), \tag{2.4}$$

where $\Omega'(\alpha, \xi)$ *is the inverse with respect to* ξ *of* $\Omega(\alpha, \xi)$, *and*

$$\omega(\alpha, \xi) = d^{\alpha}_{\xi}\alpha'.$$

Proof. By Corollary 2.2, the conclusion of Theorem 2.1 follows. From $(\alpha\beta)' = \beta'\alpha'$, $(\alpha')' = \alpha$, and the differentiability of α' and $\alpha'\beta$ in α follows the differentiability of $\alpha\beta$ in α and in β. The solvability of $\Omega(\alpha, \xi)$ in ξ follows from that of $d^{\alpha}_{\xi}\alpha'\beta$. Let $W(x, \xi) = U(x, V(x, \xi))$. Then from (2.1) and $T_\beta T_\alpha = T_{\alpha\beta}$ over $\mathfrak{D}$,

$$d^{\beta}_{\eta}(T_\beta T_\alpha x - T_{\alpha\beta}x) = W\big(T_{\alpha\beta}x, \Omega(\beta, \eta) - \Omega(\alpha\beta, d^{\beta}_{\eta}\alpha\beta)\big) = 0,$$

and from Lemma 2.3 and the transitivity of $T_{\alpha\beta}$ over $\mathfrak{D}$,

$$\Omega(\alpha\beta, d^{\beta}_{\eta}\alpha\beta) = \Omega(\beta, \eta),$$

from which (2.3) follows, and hence (2.4) on differentiating $\alpha\beta = (\beta'\alpha')'$.

THEOREM 2.4. *If*

(i) Λ_β *is closed over* Δ, Δ *with respect to* $\alpha\beta$, *differentiable over* Δ, Δ,

(ii) α' *is differentiable in* α *and such that for all* α, β

$$\alpha(\alpha'\beta) = \beta, \qquad (\alpha\beta)' = \beta'\alpha', \qquad (\alpha')' = \alpha,$$

then equations (2.3), (2.4) *hold and* β, α *are essential over* Δ, Δ *in* Λ_β, Π_α *respectively.*

Proof. From the first part of (i), $\alpha(\beta\gamma) = (\alpha\beta)\gamma$. Then from the second part of (ii), $\alpha'(\alpha\beta) = (\beta\alpha)\alpha' = (\beta\alpha')\alpha = \beta$. Hence $\alpha\beta$, $\alpha\beta'$ are mutually inverse in α : $(\alpha\beta')\beta = (\alpha\beta)\beta' = \alpha$. Similarly $\alpha\beta$, $\alpha'\beta$ are mutually inverse in β. By Lemma 2.2 and the second part of (i) $d^{\alpha}_{\xi}\alpha\beta$, $d^{\beta}_{\eta}\alpha\beta$ are solvable in ξ, η respectively, and hence β, α are essential over Δ, Δ in Λ_β, Π_α respectively. Moreover, from the first part of (ii) and the solvability of a product of solvable linear transformations, $d^{\alpha}_{\xi}\alpha'\beta = d^{\alpha'}_{\sigma\ \omega(\alpha,\xi)}\sigma\beta$ is solvable in ξ. Hence (i) and (ii) of Theorem 2.3 are satisfied, and (2.3), (2.4) hold.

COROLLARY 2.3. *If*

(i) Λ_β *is differentiable over* Δ, Δ *and generates a group over* Δ, Δ *with respect to* $\alpha\beta$, μ, α',

(ii) α' *is differentiable in* α, *then the conclusion of Theorem* 2.4 *holds.*

Proof. (i) and (ii) of Theorem 2.4 are evidently satisfied.

THEOREM 2.5. *If*

(i) *for all* α, ξ, x : $W(x, \xi)$, $\Omega(\alpha, \xi)$ *are in* $\mathfrak{D}, \Delta$ *respectively, are linear in* ξ, *and* $\Omega'(\alpha, \xi)$, *the inverse of* $\Omega(\alpha, \xi)$, *exists,*

(ii) *for all* $\alpha, \beta, \xi, \eta, x$

$$\begin{cases} d_\xi^\alpha T_a x = W(T_a x, \Omega(\alpha, \xi)) \\ T_\mu x = x, \end{cases} \tag{2.5}$$

$$\begin{cases} d_\eta^\beta \Lambda_\beta \alpha = \Omega'(\Lambda_\beta \alpha, \Omega(\beta, \eta)) \\ \Lambda_\mu \alpha = \alpha, \end{cases} \tag{2.6}$$

(iii) *the solution of* (2.5) *is unique for each* x,

then T_a *is closed over* $\Delta, \mathfrak{D}$ *with respect to* $\alpha\beta$, μ *is a unit in* Δ *over* $\mathfrak{D}$ *of* T_a, *and if*

(iv) *the solution of* (2.6) *is unique for each* α,

then T_a *generates an integral semi-group over* $\Delta, \mathfrak{D}$ *with respect to* $\alpha\beta$, μ.

Proof. Suppose (i), (ii), (iii). Then from (2.5), (2.6)

$$\begin{aligned} d_\eta^\beta T_{a\beta} x &= W\big(T_{a\beta} x, \Omega\big(\alpha\beta, \Omega'(\alpha\beta, \Omega(\beta, \eta))\big)\big) \\ &= W(T_{a\beta} x, \Omega(\beta, \eta)), \end{aligned}$$

since $\Omega'(\beta, \eta)$ is the inverse with respect to η of $\Omega(\beta, \eta)$. From (2.5)

$$d_\eta^\beta T_\beta T_a x = W(T_\beta T_a x, \Omega(\beta, \eta)).$$

Moreover $T_{a\mu} = T_\mu T_a = T_a$ over $\mathfrak{D}$. Hence by (iii) the first conclusion follows. Now suppose (iv). From (2.6)

$$\begin{aligned} d_\xi^\gamma \Lambda_{\beta\gamma} \alpha &= \Omega'\big(\Lambda_{\beta\gamma}\alpha, \Omega\big(\beta\gamma, \Omega'(\beta\gamma, \Omega(\gamma, \xi))\big)\big) \\ &= \Omega'(\Lambda_{\beta\gamma}\alpha, \Omega(\gamma, \xi)), \end{aligned}$$

and

$$d_\xi^\gamma \Lambda_\gamma \Lambda_\beta \alpha = \Omega'(\Lambda_\gamma \Lambda_\beta \alpha, \Omega(\gamma, \xi)).$$

Since also $\Lambda_{\beta\mu} = \Lambda_\mu \Lambda_\beta = \Lambda_\beta$ over Δ, (iv) implies $\Lambda_\gamma \Lambda_\beta = \Lambda_{\beta\gamma}$ over Δ, that is, $(\alpha\beta)\gamma = \alpha(\beta\gamma)$, and the last conclusion follows.

Theorems 2.1 and 2.5 together form an abstraction of Lie's first fundamental theorem.[5]

3. *The structural function* $\Gamma(\alpha, \xi, \eta)$. The definitions and notations of section 2 will apply throughout this section.

For all α, ξ, x, let $W(x, \xi)$, $\Omega(\alpha, \xi)$ be in $\mathfrak{D}, \Delta$ respectively, linear in ξ, and such that $d_z^x W(x, \xi)$, $d_\eta^\alpha \Omega(\alpha, \xi)$ exist, continuous in x and in α respectively. Then it can be shown [6] that the functions are totally differentiable in x, ξ and in α, ξ respectively. Suppose that $\Omega'(\alpha, \xi)$, the inverse with respect to ξ of $\Omega(\alpha, \xi)$, exists for all α, ξ, x. Then $\Omega'(\alpha, \xi)$ is linear [7] in ξ, differentiable [8] in α, and hence [6] totally differentiable in α, ξ.

Now if $d_z^x W(x, \xi)$ is continuous in x and T_a is transitive over $\mathfrak{D}$ and satisfies

$$d_\xi^\alpha T_a x = W(T_a x, \Omega(\alpha, \xi)) \tag{3.1}$$

for all α, ξ, x, then, since $d_\xi^\alpha d_\eta^\alpha T_a x$ is symmetric [9] in ξ, η,

$$\nabla_\eta^\xi W\big(x, \Omega(\alpha, \xi);\ W(x, \Omega(\alpha, \eta))\big) = W(x, \nabla_\eta^\xi \Omega(\alpha, \eta; \xi))$$

or

$$\nabla_\eta^\xi W(x, \eta; W(x, \xi)) = W(x, \Gamma(\alpha, \xi, \eta)), \tag{3.2}$$

where the effect of ∇_η^ξ is to subtract from its operand the result of interchanging ξ, η therein, and

$$\Gamma(\alpha, \xi, \eta) = \nabla_\eta^\xi \Omega(\alpha, \Omega'(\alpha, \xi);\ \Omega'(\alpha, \eta)) \tag{3.3}$$

is *the structural function over* Δ *of* T_a.

(3.2) is equivalent to the condition of complete integrability [10] for (3.1); this condition is therefore necessary for the existence of transitive solutions T_a of (3.1) when $W(x, \xi)$ is continuous in x, and it is also an important part of the least restrictive sufficient condition [11] known to us.

---

[5] S. Lie, *Transformationgruppen*, vol. I and vol. III.

[6] A. D. Michal, *Annali di Matematica, loc. cit.*; V. Elconin, Thesis (1937).

[7] J. Schauder, *Studia Mathematica*, vol. 2 (1930), pp. 1-6. See also S. Banach, *Studia Mathematica*, vol. 1 (1929), pp. 223-239.

[8] A. D. Michal and V. Elconin, "Completely integrable differential equations in abstract spaces." This memoir on existence theorems will appear in the *Acta Mathematica*.

[9] M. Kerner, *Annals of Mathematics*, vol. 34 (1933), pp. 546-572.

[10] A. D. Michal and V. Elconin, *Proceedings of the National Academy of Sciences*, vol. 21 (1935), pp. 534-536.

[11] A. D. Michal and V. Elconin, *Acta Mathematica, loc. cit.*

In the same way, if $d_\eta^\alpha \Omega(\alpha, \xi)$ is continuous in α and Λ_β is transitive over Δ and satisfies

$$d_\eta^\beta \Lambda_\beta \alpha = \Omega'(\Lambda_\beta \alpha, \Omega(\beta, \eta)) \tag{3.4}$$

for all α, β, η, then

$$\nabla_\eta^\xi \Omega'(\alpha, \eta; \Omega'(\alpha, \xi)) = \Omega'(\alpha, \Gamma(\beta, \xi, \eta)), \tag{3.5}$$

which is equivalent to the condition of complete integrability for (3.4); if moreover $(\alpha')' = \alpha$, $d_\xi^\alpha d_\eta^\alpha \alpha'$ exists, and

$$d_\xi^\alpha \alpha' = \omega(\alpha, \xi), \tag{3.6}$$

then, by differentiating $(\alpha')' = \alpha$: $\omega(\alpha, \xi)$, $\omega(\alpha', \xi)$ are mutual inverses with respect to ξ, and it is not difficult to show that if the conditions of complete integrability for (3.4) and (3.6) are satisfied so is that of

$$d_\xi^\alpha \alpha\beta = \omega\big((\alpha\beta)', \Omega'\big((\alpha\beta)', \Omega(\alpha', \omega(\alpha, \xi))\big)\big). \tag{3.7}$$

Several equations equivalent to (3.3) are sometimes useful; for example

$$\nabla_\eta^\xi d_\eta^\alpha \Omega(\alpha, \xi) = \Gamma(\alpha, \Omega(\alpha, \xi), \Omega(\alpha, \eta)), \tag{3.8}$$

an abstraction of Maurer's equations. Or, differentiating $\Omega(\alpha, \Omega'(\alpha, \xi)) = \xi$,

$$\Omega\big(\alpha, \nabla_\eta^\xi \Omega'(\alpha, \eta; \Omega'(\alpha, \xi))\big) = \nabla_\eta^\xi \Omega(\alpha, \Omega'(\alpha, \xi); \Omega'(\alpha, \eta)),$$

and hence

$$\Gamma(\alpha, \xi, \eta) = \Omega\big(\alpha, \nabla_\eta^\xi \Omega'(\alpha, \eta; \Omega'(\alpha, \xi))\big), \tag{3.9}$$

or

$$\nabla_\eta^\xi \Omega'(\alpha, \eta; \Omega'(\alpha, \xi)) = \Omega'(\alpha, \Gamma(\alpha, \xi, \eta)). \tag{3.10}$$

THEOREM 3.1.

(i) *The structural function* $\Gamma(\alpha, \xi, \eta)$ *is bilinear and alternating in* ξ, η.

(ii) *If* $W(x, \xi)$ *is definite in* ξ *with respect to* x *and the condition of complete integrability for* (3.1) *is satisfied, then* $\Gamma(\alpha, \xi, \eta)$ *is constant in* α.

(iii) *The condition of complete integrability for* (3.4) *is satisfied if and only if* $\Gamma(\alpha, \xi, \eta)$ *is constant in* α.

Proof. From (3.3), $\Gamma(\alpha, \xi, \eta)$ is evidently alternating in ξ, η, and additive in ξ and in η. The continuity in ξ and in η, and hence (i) follows from a theorem[12] of Banach on the limit of a sequence of additive functions.

[12] S. Banach, *Fundamenta Mathematicae*, vol. 3 (1922), pp. 133-181.

If (3.2) holds, $W(x, \Gamma(\alpha, \xi, \eta))$ is evidently constant in α, hence so is $\Gamma(\alpha, \xi, \eta)$ if $W(x, \xi)$ is definite in ξ with respect to x.

Since $\Omega(\alpha, \xi)$ is uniquely solvable in ξ, $\Omega(\alpha, \xi) = 0$ implies $\xi = 0$. Hence (iii) follows on comparing (3.5) and (3.10).

Another important property of $\Gamma(\alpha, \xi, \eta)$ can be established by means of an abstraction of Jacobi's Identity.

Let y be in S, θ in Σ, and for any function $F(x)$ on $\mathfrak{D}$ to S such that $d_y^x d_z^x F(x)$ exists, continuous in x, let

$$X_\xi F(x) = F(x; A(x, \xi))$$

where $A(x, \xi)$ is on $\mathfrak{D}, \Sigma$ to S, linear in ξ, and $d_y^x d_z^x A(x, \xi)$ exists, continuous in x.

Let

$$\begin{aligned}(X_\xi, X_\eta)F(x) &= \nabla_\eta^\xi X_\xi (X_\eta F(x)) \\ &= \nabla_\eta^\xi F(x; A(x, \xi); A(x, \eta)) + F\big(x; \nabla_\eta^\xi A(x, \eta; A(x, \xi))\big) \\ &= F\big(x; \nabla_\eta^\xi A(x, \eta; A(x, \xi))\big)\end{aligned}$$

since $d_y^x d_z^x F(x)$ is symmetric in y, z. Then Jacobi's Identity becomes

$$\Sigma((X_\xi, X_\eta), X_\theta)F(x) = 0, \tag{3.11}$$

where the summation is over the cyclical permutations of ξ, η, θ; for $((X_\xi, X_\eta), X_\theta)F(x)$ equals

$$\begin{aligned}&F\big(x; A(x, \theta); \nabla_\eta^\xi A(x, \eta; A(x, \xi))\big) \\ -&F\big(x; \nabla_\eta^\xi A(x, \eta; A(x, \xi)); A(x, \theta)\big) \\ +&F\big(x; \nabla_\eta^\xi A\big(x, \theta; A(x, \eta; A(x, \xi))\big)\big) \\ -&F\big(x; \nabla_\eta^\xi A\big(x, \eta; A(x, \xi; A(x, \theta))\big)\big) \\ -&F\big(x; \nabla_\eta^\xi A\big(x, \eta; A(x, \xi); A(x, \theta)\big)\big),\end{aligned}$$

in which the first and second terms cancel, the sums in (3.11) arising from the third and fourth terms cancel, and the sum arising from the fifth term equals zero.

THEOREM 3.2. *If*

(i) *$W(x, \xi)$ is definite in ξ with respect to x and*

(ii) *for all x, y, z, ξ $d_y^x d_z^x W(x, \xi)$ exists, continuous in x, or*

(iii) *for all $\alpha, \eta, \theta, \xi$ $d_\eta^\alpha d_\theta^\alpha \Omega'(\alpha, \xi)$ exists, continuous in α, then*

$$\sum_{\xi, \eta, \theta} \Gamma(\alpha, \Gamma(\alpha, \xi, \eta), \theta) = 0, \tag{3.12}$$

where the summation is cyclical.

Proof. Suppose (i), (ii), and $A(x,\xi) = W(x,\xi)$ for all x, ξ. Then by (3.2)

$$(X_\xi, X_\eta)F(x) = F\big(x; \nabla_\eta^\xi W(x,\eta; W(x,\xi))\big)$$
$$= F\big(x; W(x, \Gamma(\alpha,\xi,\eta))\big) = X_{\Gamma(\alpha,\xi,\eta)}F(x),$$

and hence

$$\sum_{\xi,\eta,\theta} ((X_\xi, X_\eta), X_\theta)F(x) = X_J F(x) = 0$$

by Jacobi's Identity, where J is the left member of (3.12). If in particular $F(x) = x$, for all x, then $X_J F(x) = W(x, J) = 0$, and from (i), $J = 0$.

Now suppose (iii), S is $\mathfrak{Z}$, $\mathfrak{D}$ is Δ, and $A(\alpha,\xi) = \Omega'(\alpha,\xi)$ for all α, ξ. Then by (3.10) and Jacobi's Identity, with $F(x) = x$: $\Omega'(\alpha, J) = 0$. Hence $J = 0$, since $\Omega'(\alpha,\xi)$ is uniquely solvable in ξ.

4. *Illustrative groups.* The results in the preceding sections can be illustrated by a group of matrix transformations.

Let $S, \mathfrak{Z}$ be respectively the set of all matrices with m rows, n columns, and the set of all matrices with n rows, n columns, where in each set the matrix elements are in the field of all complex numbers. For any $\omega = (\omega_{ij})$ in S or in $\mathfrak{Z}$, let $\| \omega \| = \max | \omega_{ij} |$. When $\| \omega \|$ as the norm of ω, and the usual definitions of matrix addition and multiplication, $S, \mathfrak{Z}$ are Banach spaces, and $x\alpha$, $\alpha\beta$ are bilinear functions on $S, \mathfrak{Z}$ to S and on $\mathfrak{Z}, \mathfrak{Z}$ to $\mathfrak{Z}$ respectively.

Let $\mathfrak{D}$ be S, Δ the set of non singular matrices in $\mathfrak{Z}$, μ the identity matrix in $\mathfrak{Z}$, α' the inverse matrix of α for all α in Δ, and let $T_\alpha x = x\alpha$ for all α, x in $\Delta, \mathfrak{D}$. Clearly Δ is a domain. For brevity, let α, β be in Δ; ξ in $\mathfrak{Z}$; x, z in $\mathfrak{D}$.

T_α generates a group over $\Delta, \mathfrak{D}$ with respect to $\alpha\beta$, μ, α'.

T_α is differentiable over $\Delta, \mathfrak{D}$, since for all α, x, ξ, z

$$d_{\xi z}^{\alpha x} T_\alpha x = x\xi + z\alpha.$$

$d_\xi^\alpha \alpha'$ exists equal to $-\alpha'\xi\alpha'$. This follows from a known theorem,[13] on noting that $\alpha\xi$, $\alpha'\xi$ are mutually inverse with respect to ξ. Or it may be deduced directly as follows. $(\alpha+\xi)' - \alpha' = -(\alpha+\xi)'\xi\alpha'$ if $\| \xi \|$ is sufficiently small. From the definition of matrix multiplication, $\| \alpha\beta \| \leq n \| \alpha \| \| \beta \|$, and, say from Kramer's Rule, $\|(\alpha+\xi)'\|$ is bounded for all sufficiently small $\| \xi \|$. Hence

$$\| (\alpha+\xi)' - \alpha' \| \leq n^2 \| (\alpha+\xi)' \| \| \xi \| \| \alpha' \|$$

[13] A. D. Michal and V. Elconin, *Acta Mathematica, loc. cit.* This theorem asserts the differentiability with respect to a parameter of the inverse of a solvable linear function itself differentiable in the parameter.

is arbitrarily small for all sufficiently small $\| \xi \|$: α' is continuous in α; and for any positive number ϵ

$$\| (\alpha + \xi)' - \alpha' + \alpha' \xi \alpha' \| = \| \{ (\alpha + \xi)' - \alpha' \} \xi \alpha' \| \\ \leq n^2 \| (\alpha + \xi)' - \alpha' \| \ \| \alpha' \| \ \| \xi \| \leq \epsilon \| \xi \|$$

if $\| \xi \|$ is sufficiently small. Since $-\alpha' \xi \alpha'$ is also linear in ξ, $d^{\alpha}_{\xi} \alpha' = -\alpha' \xi \alpha'$.

With the notations introduced in the proof of Theorem 2. 1,

$$U(x, z) = -z\delta', \qquad V(x, \xi) = x\xi, \qquad \Omega(\alpha, \xi) = -\alpha' \xi \delta,$$

since $\beta / \alpha = \alpha' \beta$ here. Hence by that theorem

$$d^{\alpha}_{\xi} T_{\alpha} x = -(T_{\alpha} x)(-\alpha' \xi \delta) \delta' = (T_{\alpha} x) \alpha' \xi, \tag{4. 1}$$

which may be directly verified.

For simplicity, let $\delta = -\mu$. With the notations of Theorem 2. 3,

$$\Omega'(\alpha, \xi) = \alpha \xi, \qquad \omega(\alpha, \xi) = -\alpha' \xi \alpha'.$$

Hence by that theorem

$$d^{\beta}_{\eta} \alpha \beta = (\alpha \beta)(\beta' \eta) = \alpha \eta, \tag{4. 2}$$

and

$$d^{\alpha}_{\xi} \alpha \beta = -(\alpha \beta) \{ (\alpha \beta)' \alpha (-\alpha' \xi \alpha') \} (\alpha \beta) = \xi \beta, \tag{4. 3}$$

which are evidently true.

With the notations of section 3,

$$\Gamma(\alpha, \xi, \eta) = \alpha' \nabla^{\xi}_{\eta} \alpha \eta \xi = \eta \xi - \xi \eta,$$

which is constant in α, and hence the condition of complete integrability for (4. 2) is satisfied. This is already evident since the transformation Λ_{β} defined by $\Lambda_{\beta} \alpha = \alpha \beta$ is transitive over Δ. For similar reasons, (4. 1) and (4. 3) are completely integrable. Moreover, the solutions of these equations are uniquely determined by one-point initial conditions. Thus T_{α} satisfies

$$\begin{cases} d^{\alpha}_{\xi} T_{\alpha} x = (T_{\alpha} x) \alpha' \xi \\ T_{\mu} x = x. \end{cases} \tag{4. 4}$$

Conversely, let $T_{\alpha} x$ now be any solution of (4. 4). Then

$$d^{\alpha}_{\xi} (T_{\alpha} x) \alpha' = (d^{\alpha}_{\xi} T_{\alpha} x) \alpha' - (T_{\alpha} x)(\alpha' \xi \alpha') = ((T_{\alpha} x) \alpha' \xi) \alpha' - (T_{\alpha} x)(\alpha' \xi \alpha') = 0.$$

Hence for some y in $\mathfrak{D}$, $(T_{\alpha} x) \alpha' = y$, or $T_{\alpha} x = y a$ for all α, and from $T_{\mu} x = x$ follows $x = y$. Essentially the same device may be used on (4. 2), (4. 3).

The above results all hold when the matrix transformation group is replaced by a certain group of linear functional transformations of the third kind.[14] Let S be the well-known Banach space of functions x^σ, real and continuous over $a \leqq \sigma \leqq b$, with $\| x^\sigma \| = \max_{a \leq \sigma \leq b} | x^\sigma |$. Let Σ be the set of all ordered pairs $(\alpha^\sigma, \alpha^\sigma_\tau)$ of functions α^σ and α^σ_τ, real and continuous for $a \leq \sigma, \tau \leq b$. Let

$$\| (\alpha^\sigma, \alpha^\sigma_\tau) \| = \max_{a \leq \sigma \leq b} | \alpha^\sigma | + \max_{a \leq \sigma, \tau \leq b} | \alpha^\sigma_\tau |$$

$$(\alpha^\sigma, \alpha^\sigma_\tau) + (\beta^\sigma, \beta^\sigma_\tau) = (\alpha^\sigma + \beta^\sigma, \alpha^\sigma_\tau + \beta^\sigma_\tau),$$

and

$$r(\alpha^\sigma, \alpha^\sigma_\tau) = (r\alpha^\sigma, r\alpha^\sigma_\tau),$$

for all $(\alpha^\sigma, \alpha^\sigma_\tau)$, $(\beta^\sigma, \beta^\sigma_\tau)$ in Σ and all real numbers r. Then Σ is a Banach space.[15] Let x, α, β be x^σ, $(\alpha^\sigma, \alpha^\sigma_\tau)$, $(\beta^\sigma, \beta^\sigma_\tau)$ respectively, and define

$$x\alpha = x^\sigma \alpha^\sigma + x^\rho \alpha^\sigma_\rho,$$

$$\alpha\beta = (\alpha^\sigma \beta^\sigma, \alpha^\sigma_\tau \beta^\sigma + \alpha^{(\tau)} \beta^\sigma_\tau + \alpha^\rho_\tau \beta^\sigma_\rho),$$

where in any term an index occurring once as subscript, once as superscript, and not in parentheses, indicates Riemann integration over (a, b) with respect to that index; thus $\alpha^\rho_\tau \beta^\sigma_\rho = \int_a^b \alpha^\rho_\tau \beta^\sigma_\rho d\rho$, but τ is free in $\alpha^{(\tau)} \beta^\sigma_\tau$. $x\alpha$, $\alpha\beta$ are bilinear functions on S, Σ to S and on Σ, Σ to Σ respectively. Let $T_a x = x\alpha$. Then T_a, a linear functional transformation of the third kind, generates a semi-group over Σ, S with respect to $\alpha\beta$. Let $\mathfrak{D}$ be S, Δ the set of all α in Σ for which $\alpha^\tau \neq 0$ and the Fredholm determinant $D[\alpha^\sigma_\tau / \alpha^{(\tau)}] \neq 0$, Γ^σ_τ the resolvent kernel of $\alpha^\sigma_\tau / \alpha^{(\tau)}$, and let $\mu^\sigma = 1$, $\mu^\sigma_\tau = 0$ for $a \leq \sigma, \tau \leq b$. Then[16] T_a generates a group over Δ, $\mathfrak{D}$ with respect to $\alpha\beta$, μ, α', where $\mu = (\mu^\sigma, \mu^\sigma_\tau)$, $\alpha' = (1/\alpha^\sigma, \Gamma^\sigma_\tau / \alpha^{(\sigma)})$. The results obtained for the matrix transformation group, and the arguments leading to them, except that employed in the direct deduction of $d^a_\xi \alpha' = -\alpha' \xi \alpha'$, all hold with the new interpretations of $T_a x$, $\alpha\beta$.

5. *Conclusion.* The main results of this paper express differential properties of abstract differentiable transformation groups and certain of their generalizations. The study of the intimate group structure, which we have

---

[14] A. D. Michal and T. S. Peterson, *Annals of Mathematics*, vol. 32 (1931), pp. 431-450. See also A. D. Michal, *American Journal of Mathematics*, vol. 50 (1928), pp. 473-517.

[15] A. D. Michal and R. S. Martin, *Journal de Mathématiques Pures et Appl.*, vol. 13 (1934), pp. 69-91.

[16] A. D. Michal and T. S. Peterson, *loc. cit.*

already begun, suggests that the totality of non-trivial differential properties is very large. It is at least reasonable to expect that salient formal group properties strongly imply and may sometimes be characterized by differential properties. The determinations of analytic conditions under which the latter properties imply the former constitute the most difficult and laborious problems in the theory. For example, in Theorem 2.5, it was shown that the uniqueness of the solutions of certain differential equations implied fundamental formal group properties of these solutions, but the determination of reasonably broad analytic conditions which insure this uniqueness is difficult; our best results so far seem too restrictive. It is probable, however, that better conditions will be obtained by strengthening and extending basic existence theorems for completely integrable abstract differential equations which we have already established.[17]

CALIFORNIA INSTITUTE OF TECHNOLOGY,
PASADENA, CALIFORNIA.

[17] A. D. Michal and V. Elconin, *Acta Mathematica, loc. cit.*

REMARKS ON SYSTEMS OF ORDINARY DIFFERENTIAL EQUATIONS.[1]

By E. R. VAN KAMPEN.

I. *A uniqueness theorem.* Let the vector $\phi = \phi(\xi) = (f_1, \cdots, f_n)$ be a continuous function of the vector $\xi = (x_1, \cdots, x_n)$ in a certain region R of the ξ-space. Let t (or s) be a real variable and let a subscript t (or s) denote total differentiation with respect to this variable. It is well known that if ξ_0 is a fixed vector in R, then the system of ordinary differential equations

$$\xi_t = \phi(\xi) \tag{1}$$

has at least one solution

$$\xi = \psi(\xi_0, t), \text{ for which } \psi(\xi_0, 0) = \xi_0 \tag{2}$$

and which is defined for all t in a certain interval which contains $t = 0$ and depends on ξ_0. Several conditions concerning the function $\phi(\xi)$ are known, each of which implies that the solution (2) of (1) is uniquely determined by ξ_0. In the uniqueness theorem [2] stated below, the same conclusion is drawn from an assumption on a certain set of solutions of (1), without any further restriction on the function $\phi(\xi)$ itself. Using the symbol $|\eta|$ to denote the length of the vector η, this uniqueness theorem may be formulated as follows:

Let the equation (1) *have a solution* (2) *such that*

(i) *one and only one curve of the system* (2) *passes through any point of* R, i. e.,

$$\psi(\psi(\xi_0, t), s) = \psi(\xi_0, t + s), \tag{3}$$

whenever the expression on the left side is defined and

(ii) *the function* (2) *satisfies a Lipschitz condition with respect to* ξ_0, i. e., *there exsts a positive number* C *which is independent of* ξ_0 *and* η_0 *such that*

[1] Received November 20, 1936.

[2] This theorem combined with a treatment of Lie groups, obtained by the author several years ago but never published, implies that a group nucleus homeomorphic with an n-cell is a Lie group nucleus if (i) the functions $c\gamma = f\gamma(a\alpha, b\beta)$ defining the group operation have a continuous derivative with respect to $a\alpha$ at a fixed value of $a\alpha$ and (ii) the functions $f\gamma(a\alpha, b\beta)$ satisfy a Lipschitz condition with respect to $b\beta$. A publication of the considerations involved is planned in the near future.

$$|\psi(\xi_0, t) - \psi(\eta_0, t)| < C\,|\xi_0 - \eta_0| \tag{4}$$

for any ξ_0 and η_0 in R.

Then, if ξ_0 is a fixed vector in R, (2) defines the only solution of (1) which reduces for $t = 0$ to ξ_0.

Let $\chi(t)$, $t_1 < t < t_2$, be any fixed solution of (1) and put

$$\zeta(s, t) = \psi(\chi(s), t - s), \tag{5}$$

where $t_1 < s < t_2$, while, for a given s, the variable t is restricted in such a way that the right side of (5) has a meaning. Obviously (5) determines, for a fixed s, the unique curve of the system (2) through the point $\chi(s)$, so that the theorem is equivalent to the statement that $\zeta(s, t)$ is independent of s, whenever $\chi(s)$ is a solution of (1).

First, it is seen from (2) and (5) that $\zeta(t, t) = \chi(t)$, so that

$$\zeta_t(t, t) = \phi(\zeta(t, t)), \tag{6}$$

since $\chi(t)$ is a solution of (1). Also, since (2) is a solution of (1), it follows from (5) that $\zeta_t(s, t) = \phi(\zeta(s, t))$. Thus one obtains, on placing $t = s$ in the last equation,

$$\{\zeta_t(s, t)\}_{t=s} = \phi(\zeta(s, s)). \tag{7}$$

On placing $t = s$ in (6) and subtracting the result from (7), one obtains

$$\{[\zeta(s, t) - \zeta(t, t)]_t\}_{t=s} = 0. \tag{8}$$

Hence, if $\epsilon > 0$ is given and s is fixed so that $t_1 < s < t_2$, finally if s' is sufficiently near s, it follows from $\{\zeta(s, t) - \zeta(t, t)\}_{t=s} = 0$ that

$$|\zeta(s, s') - \zeta(s', s')| < \epsilon\,|s - s'|. \tag{9}$$

On the other hand, if s and t are fixed in such a way that $\zeta(s, t)$ is defined, and s' is sufficiently near s, then $\zeta(s', t)$ is defined, so that (3) and (5) imply that

$$\zeta(s, t) = \psi(\zeta(s, s'), t - s'), \qquad \zeta(s', t) = \psi(\zeta(s', s'), t - s'). \tag{10}$$

On using condition (ii) in the form (4), one obtains from (10)

$$|\zeta(s, t) - \zeta(s', t)| < C\,|\zeta(s, s') - \zeta(s', s')|. \tag{11}$$

From (9) and (11) it is clear that

$$|\zeta(s,t)-\zeta(s',t)|<\epsilon C\,|s-s'| \tag{12}$$

whenever $\zeta(s,t)$ is defined and s' is sufficiently near s. Since C is independent of s and s', it is seen from (12) that $\zeta_s(s,t)$ exists and is equal to 0 whenever $\zeta(s,t)$ is defined. Thus $\zeta(s,t)$ is independent of s, so that the theorem is proved.

Needless to say, the solution of (1) need not be unique if a system (2) of solutions is known satisfying (i) but not satisfying (ii).

II. *The set of split points of an ordinary differential equation on a solution.* If the vector function $\phi(\xi,t)$ of the vector ξ and the real variable t is continuous in a region R of the (ξ,t)-space, then the system

$$\xi_t=\phi(\xi,t) \tag{1}$$

may have more than one solution passing through a given point of R. However, very little is known about the possible distribution in R of those points through which passes more than one solution of the system. The present note contains a characterization of the distribution of such points on a solution of the single equation

$$x_t=f(x,t), \tag{2}$$

where $f(x,t)$ is a continuous function in a region R of the (x,t)-plane, and also contains a remark on the corresponding problem in case of the system (1).

It is convenient to order the points on a solution $x=g(t)$ of (1) by saying that the point P_1 is to the right of the point P_2 if P_1 belongs to a larger value of t than P_2. Then it is clear what is meant by the following definitions: The point $P\equiv(x,t)$ is called a *right split point* (a *left split point*) of (2) if two solutions of (2) exist, both containing P and such that no arc of the first solution with P as left (right) end point belongs entirely to the second solution. A point set S on a solution of (1) is said to be *right* (*left*) *closed* if any point of R which is the limit of a sequence of points on S belongs to S, provided each point of the sequence is to the right (left) of the next point. It will be proved first that:

(3) *The set of right (left) split points on any solution of* (2) *is right* (*left*) *closed.* In fact, it is well known that through any point P of R there passes a uniquely determined right upper, right lower, left upper, left lower solution of (2). It is clear that if P_1 and P_2 are on a solution α of (1) and

P_1 is to the right of P_2, then the arc P_2P_1 of α together with the right upper solution through P_1, is a right hand solution through P_2, so that the right upper solution through P_2 is never below the right hand upper solution through P_1. Similar statements hold for the other extreme solutions. Now let $\{P_n\}$ be a sequence of right split points on a solution α of (2) such that P_n is to the right of P_{n+1}, for every n, and $\{P_n\}$ converges to a point P of α. Then the right upper (lower) solution of (2) through P is never lower (higher) than the right upper (lower) solution through P_n. Thus the right hand upper and lower solutions through P_n do not coincide on any segment with P as endpoint, since the right hand upper and lower solutions through the right split point P_n do not coincide on any segment with P_n as endpoint, and any segment of α with P as left end point contains at least one point P_n. The case of left split points may be similarly treated.

(4) *If on the t-axis in the (x, t)-plane there are given point sets S_l and S_r which are left and right closed respectively, then there exists a continuous function $f(x, t)$, such that $f(0, t) = 0$ and S_l, S_r are the sets of left, right split points of* (2) *on the t-axis respectively.*

The function $f(x, t)$ will be constructed in such a way that $f(x, t) \geqq 0$ for any x and t. Then the t-axis is the right lower and left upper solution through any of its points, so that the values of $f(x, t)$ in the lower (upper) half plane alone are sufficient to decide whether any point of the t-axis is a left (right) split point or not. Thus it will be sufficient to consider the case where S_l is the empty set and $f(x, t) = 0$ whenever $x \leqq 0$ and for all t, while $f(x, t) \geqq 0$ whenever $x > 0$ and for all t.

Since the set S_r, which will from now on be denoted by S, is right closed by assumption, so that every point not in S is the left end point of an interval that does not meet S, the complement C of S on the t-axis is the sum of an at most enumerable collection of mutually disjoint intervals, the left end point of each of which is in S, while the right end point may or may not be in S. The set S itself may but need not contain intervals. In any case the set of maximal intervals contained in S is at most enumerable, and the left end point of any such interval is always contained in S. Let an enumerable set T contain the following points:

(i) The left end points of all maximal intervals in S;

(ii) The end points of all intervals constituting C;

Then any point s of S has one of the following properties:

(iii) s is contained in T;

(iv) s is interior point of a maximal interval of S.

(v) s is the limit of a monotone decreasing sequence of elements both in S and in T.

In fact, if an element s of S does not have either property (iii) or property (iv), then s is not the left end point of any interval belonging entirely to S or entirely to C. But then any interval of which s is the left end point contains the left end point of one of the intervals constituting C. Since such a left end point is contained both in S and in T, the point s has property (v).

Now let the points of T be arranged into the sequence $\{t_n\}$. Suppose that a function $f(x, t)$ has been defined on closed sets $A_1, \cdots, A_{n-1}$, corresponding to the points $t_1, \cdots, t_{n-1}$ and let there be defined, corresponding to the point t_n of T, a closed set A_n and a function $f(x, t)$ on A_n as follows:

(vi) If t_n is not the left end point of an interval either in S or in C, let A_n be the curve

$$(*) \qquad x = (t - t_n)^2/n, \qquad t_n \leqq t \leqq t_n + \tau_n,$$

where either $\tau_n = 1$ or, if the smallest number $p > 0$ such that the point $(p^2/n,\ t_n + p)$ is on one of the sets $A_1, \cdots, A_{n-1}$ exists and is not larger than 1, then $\tau_n = (n-1)p/n$. The value of $f(x, t)$ at any point of A_n is defined to be the slope of (*) at that point.

(vii) If t_n is in S and is the left end point of an interval $t_n < t < r_n$ which is contained in C, let A_n be the set defined by the inequalities

$$(**) \qquad t_n \leqq t \leqq r;\ 0 \leqq x \leqq \min\,[(t - t_n)^2/n,\ 1/n],$$

while $f(x, t) = 2x/(t - t_n)$ for any point (x, t) in A_n (in particular $f(0, t_n) = 0$). Thus through any point (x, t) of A_n there passes a solution of (2) of the form $x = c(t - t_n)^2$, where $0 \leqq c \leqq 1/n$ in view of (**).

(viii) If t_n is the left end point of an interval $t_n \leqq t < r_n$ which is contained in S, let A_n be the set defined by (6), and put $f(x, t) = 2x^{\frac{1}{2}}n^{-\frac{1}{2}}$ for any point (x, t) in A_n. Thus through any point of A_n there passes a solution of (2) of the form $x = (t - c)^2/n$, where obviously $t_n \leqq c \leqq r_n$.

(ix) If t_n is in C, so that it is the left end point of an interval $t_n \leqq t < r_n$ which is contained in C, let ϵ_n be such that $0 < \epsilon_n < 1/n$, while the set A_n defined by

$$t_n \leqq t \leqq r_n;\ 0 \leqq x \leqq \min\,[(t - t_n)/\epsilon_n,\ 1/n]$$

does not have a point in common with any of the sets $A_1, \cdots, A_{n-1}$ except maybe the points $(0, t_n)$ or $(0, r_n)$. Put $f(x, t) = 0$ for any point (x, t) in A_n.

Thus there may be defined by a process of complete induction a set A_n corresponding to every element t_n of T and a continuous function $f(x, t)$ on A_n. If (x, t) is any point of A_n, then clearly $x \leqq 1/n$ and $f(t, x) \leqq 2x^{\frac{1}{2}}n^{-\frac{1}{2}}$. Since the different sets A_n may have only points on the t-axis in common, it follows that, by putting $f(x, t) = 0$ for any point on the t-axis where $f(x, t)$ is not yet defined, the function $f(x, t)$ has now been defined as a continuous function on a closed set. Put $f(x, t) = 0$ for $x \leqq 0$ and for all t and let $f(x, t)$ be defined for $x \geqq 0$ as a non-negative continuous extension [3] of the function already defined. It will be shown that a point on the t-axis is a right split point of (2) if and only if the point is in S, thus completing the proof of (4) in view of the remarks immediately following (4).

Let s be any point of S. If s satisfies (iii), then it is clear from (vi), (vii), (viii) that s is a right split point of (2). Hence, if s satisfies (v), then s is a right split point of (2) in view of (3). If s satisfies (iv), then s is a split point of (2) by (viii) in view of (i). Thus any point of S is a right split point of (2). Let r be a point of C. If r is not an interior point of C, then r is not a right split point of (2) by (ix). Finally, if r is an interior point of C, then r is not a right split point of (2) by (vii) or (ix).

Let it be remarked that the notion of a split point is not as simple in case of the system (1) as in case of the equation (2). In fact it is possible to construct a function $\phi(\xi, t)$ such that a given point $P \equiv (\xi_0, t_0)$ in the (ξ, t)-space has the following properties

Any two solutions of (1) through P are identical on a t-interval containing P_0 and

If any t-interval I containing t_0 is given, then two solutions of (1) may be found which are not identical on I.

III. *On the set of split points of an ordinary differential equation on a cross line.* Let $f(x, t)$ be a continuous function in a region U of the (x, t)-plane. A simple arc α will be said to be a cross line of the differential equation

$$x_t = f(x, t) \tag{1}$$

if any solution of (1) through a point P of α has only the point P in common with α. For a sufficiently small $\epsilon > 0$, let $R_\epsilon(L_\epsilon)$ denote the set of those points

[3] It is easily seen how in the course of this extension eventual unbounded intervals contained in S or in C must be taken in account.

of U which are on solutions through points (x_0, t_0) of α and correspond to values of t satisfying $t_0 < t < t_0 + \epsilon (t_0 - \epsilon < t < t_0)$. First it will be shown that:

(2) *If, for a fixed $\epsilon > 0$, the set R_ϵ does not contain a left split point of* (1), *then the set of right split points of* (1) *on α is at most enumerable.* In fact, if P is a right split point of (1) on α, then the right upper and right lower solutions of (1) through P in R_ϵ are distinct, so that a certain open subset U_P of R_ϵ is between these two solutions. Moreover, if P and Q are distinct right split points of (1) on α, then the open sets V_p, V_q cannot have a point in common, since (1) does not have a left split point in R_ϵ, so that solutions through distinct points of α cannot have a point in common on R_ϵ. Thus (2) follows from the well-known fact that a collection of mutually disjoint open subsets of the (x, t)-plane is at most enumerable. Note that the statement of (2) remains correct if the notion "left split point" is replaced by "right split point," even if R_ϵ is not replaced by L_ϵ.

Although (2) is true, it is quite easy to construct a function $f(x, t)$, such that (1) has cross lines consisting entirely of split points. In fact it will be shown that:

(3) *If C is any nowhere dense closed set on the t-axis, then a function $f(x, t)$ exists such that both the set of all right split points and the set of all left split points of* (1) *contains all points in the (x, t)-plane of which the t-coördinate is in C.* It follows of course from (2) that the set of split points must contain more points than postulated in (3). However, the function $f(x, t)$ will be constructed in such a way that the set of split points not postulated in (3) is enumerable. It is to be noted that, according to (3), the measure of the set of split points of (1) in a bounded region may be any non-negative number less than the measure of that region. On the other hand no example is known of a function (1) for which the set of split points is dense in any region of the (x, t)-plane. The construction of the example announced in (3) will be given in several steps:

a. Functions $f(x, t)$ may be constructed such that:

(i) $f(x, t)$ is defined and continuous for every (x, t);

(ii) $f(x, t) = 0$ if either $x \leqq 0$ or $x \geqq 1$;

(iii) $|f(x, t)| < M$, where M is independent of x and t;

(iv) Any solution of (1) which is defined for all t contains at least one right split point and at least one left split point.

(v) The set of split points of (1) is enumerable.

Such a function may, for instance, be defined as follows: At any point P of any one of the curves

$$x = 4n + c(1 - \cos 2\pi t) \quad \text{or} \quad x = 4n + 2 + c(1 + \cos 2\pi t), \quad 0 \leqq t \leqq 1,$$

where c is any real number such that $-1 \leqq c \leqq 1$ and n is any positive or negative integer, put $f(x, t)$ equal to the slope at P of the curve passing through P.

b. Functions $f(x, t)$ may be constructed which have properties (i), (ii), (iii) and are such that

(vi) The set of left split points and the set of right split points on any solution of (1) which is defined for all t has the points corresponding to $t = 0$ and to $t = 1$ as cluster points;

(vii) Although, according to (ii), (vi) and (3) of Section II, every point $(x, 0)$ is a right split point of (1) and every point $(x, 1)$ is a left split point of (1), the set of split points for which $0 < t < 1$ is enumerable.

Such a function may, for instance, be defined as follows:

If $t \leqq 0$ or $t \geqq 0$, put $f(x, t) = 0$;

if $2^{-n} \leqq t \leqq 2^{1-n}$, put $f(x, t) = n^{-1} g(n2^n x, 2^n t - 1)$;

finally, if $1 - 2^{1-n} \leqq t \leqq 1 - 2^{-n}$, put $f(x, t) = n^{-1} g(n2^n x, 2^n(t - 1) + 2)$.

Here $g(x, t)$ is a function with properties (i), (ii), (iii), (iv) and (v) and n is any integer larger than one. That the function $f(x, t)$ thus defined has the required properties may be easily verified. A function with properties (i), (ii), (iii), (vi) and (vii) will be denoted by $h(x, t)$.

c. Now let C be a given nowhere dense closed set on the t-axis, and let the intervals constituting the complement of C be [4] $a_n < t < b_n$, $n = 1, 2, \cdots$.

A function $f(x, t)$, satisfying the requirements of (3) may now be defined as follows:

[4] It is easy to see how an eventual unbounded interval in C should be taken in account.

If t is in C, put $f(x, t) = 0$;

if $a_n \leqq t \leqq b_n$, for some n, put $f(x, t) = \frac{1}{n} h\left(\frac{nx}{b_n - a_n}, \frac{t - a_n}{b_n - a_n}\right)$.

The function $f(x, t)$ thus defined clearly is continuous for every x and t, since $h(x, t)$ has properties (i), (ii), (iii). Let t_0 be an element of C. Since C is nowhere dense, either there exists an n such that $t_0 = a_n < b_n$ or, for every $\epsilon > 0$, there exists an n such that $t_0 < a_n < b_n < t_0 + \epsilon$. Thus it follows from (vi) and from (3) of Section II that a point $P \equiv (x_0, t_0)$ is a left split point of (1) if t_0 is in C. It is similarly seen that P also is a right split point of (1). Finally, it is clear from (vii) that the set of those split points of (1) whose t-coördinate is not in C is an enumerable set.

THE JOHNS HOPKINS UNIVERSITY.

ON A SYMMETRICAL CANONICAL REDUCTION OF THE PROBLEM OF THREE BODIES.[1]

By E. R. VAN KAMPEN and AUREL WINTNER.

Introduction. It is known [2] that the classical first integrals of the problem of three bodies permit reducing the degree of freedom from 9 to 4 in a canonical form and that on the basis of these integrals no further reduction of the general problem is possible, if the independent variable of the reduced system is the same as that of the original system, i. e., if no use is made of the energy integral. In the particular case of the planar problem of three bodies the reduced degree of freedom is 3 instead of 4. In the latter case Murnaghan [3] recently gave a symmetric treatment of the problem by using the 3 mutual distances as the 3 coördinates of the reduced system.

The object of this paper is to extend this symmetrical reduction to the non-planar case by using as the 4 coördinates the 3 mutual distances and an angular variable which is symmetrical with respect to the three bodies and becomes an ignorable coördinate in the planar case. Correspondingly, the treatment given by Murnaghan [4] for the planar case is based on a "modification," in the sense of Routh, of the Lagrangian function of the original problem. The treatment to be given in the present paper both for the planar and non-planar case is based on the consideration of the Hamiltonian instead of the Lagrangian equations. Actually, the reduction will be composed of two canonical transformations (Sections II and III) which are independent of the form of the Hamiltonian function. On applying these canonical transformations to the Hamiltonian system of the problem of three bodies (Sections IV and V), the desired reduced equations immediately follow.

It is suggested by the simplicity and symmetry of the result that the reduced equations can be applied with success to several classical problems concerning particular solutions of a preassigned nature.

I. *Extended canonical transformations and invariant relations.* The

[1] Received November 9, 1936.

[2] For the extensive literature of the subject, cf. R. Marcolongo, *Il problema dei tre corpi*, Milano, 1919, Chap. 2, pp. 27-41.

[3] F. D. Murnaghan, "A symmetric reduction of the planar three-body problem," *American Journal of Mathematics*, vol. 58 (1936), pp. 829-832.

[4] *Loc. cit.*, [3], p. 831.

object of this Section I is to collect for later use some rules dealing with the introduction of new variables into a canonical system

$$(1) \qquad q'_l = \partial H/\partial p_l, \qquad p'_l = -\partial H/\partial q_l; \qquad (l = 1, 2, \cdots, n)$$

with n degrees of freedom, where the prime denotes differentiation with respect to t and the Hamiltonian function

$$(2) \qquad H = H(q_1, \cdots, q_n, p_1, \cdots, p_n)$$

is supposed to have continuous partial derivatives of the second order in the $2n$-dimensional region under consideration.

(i) Let κ_λ and π_λ, where $\lambda = 1, \cdots, \nu$, be the coördinates and impulses in a 2ν-dimensional phase-space, where ν is less than the degree of freedom, n, of (1). Put

$$(3_1) \qquad \kappa_\lambda = \phi_\lambda(q_1, \cdots, q_n); \qquad (\lambda = 1, 2, \cdots, \nu),$$

$$(3_2) \qquad p_l = \sum_{\lambda=1}^{\nu} \pi_\lambda \, \partial\phi_\lambda/\partial q_l; \qquad (l = 1, 2, \cdots, n),$$

and restrict the n coördinates q_l by $n - \nu$ relations

$$(4) \qquad \psi_\mu(q_1, \cdots, q_n) = 0; \qquad (\mu = 1, 2, \cdots, n - \nu).$$

It will be asssumed that the n functions $\phi_1, \cdots, \phi_\nu, \psi_1, \cdots, \psi_{n-\nu}$ of the n variables $q_1, \cdots, q_n$ have in the n-dimensional q-region under consideration continuous partial derivatives of the third order and a non-vanishing Jacobian. Thus from (3_1) and (4)

$$(5) \qquad q_l = f_l(\kappa_1, \cdots, \kappa_\nu); \qquad (l = 1, 2, \cdots, n),$$

where the functions f_l are locally univalued and have continuous partial derivatives of the third order. On substituting (3_2) and (5) into (2), one obtains a function of the κ_λ and π_λ which will be denoted by H, so that, in virtue of (3_2) and (5),

$$(6) \qquad H = \mathrm{H}, \text{ where } \mathrm{H} = \mathrm{H}(\kappa_1, \cdots, \kappa_\nu, \pi_1, \cdots, \pi_\nu).$$

Now the function H thus defined has the following property: For those solutions of (1) along which (4) is satisfied the canonical system (1) is equivalent to the canonical system

$$(7) \qquad \kappa'_\lambda = \partial \mathrm{H}/\partial \pi_\lambda, \qquad \pi'_\lambda = -\partial \mathrm{H}/\partial \kappa_\lambda; \qquad (\lambda = 1, 2, \cdots, \nu)$$

of ν $(< n)$ degrees of freedom. The proof requires but straightforward differentiations and substitutions.

(ii) Let again κ_λ and π_λ, where $\lambda = 1, 2, \cdots, \nu$ and $\nu < n$, be the coördinates and impulses in a 2ν-dimensional phase-space, but replace (3_1), (3_2) by

$$(8_1) \qquad q_l = F_l(\kappa_1, \cdots, \kappa_\nu); \qquad (l = 1, 2, \cdots, n),$$

$$(8_2) \qquad \pi_\lambda = \sum_{l=1}^{n} p_l \, \partial F_l / \partial \kappa_\lambda; \qquad (\lambda = 1, 2, \cdots, \nu),$$

and (4) by

$$(9) \qquad \sum_{l=1}^{n} p_l G_{l\mu}(\kappa_1, \cdots, \kappa_\nu) = 0; \qquad (\mu = 1, 2, \cdots, n - \nu).$$

It will be assumed that, in the ν-dimensional κ-region under consideration, the n functions F_l and the $n(n - \nu)$ functions $G_{l\mu}$ have continuous partial derivatives of the third order and that, in the same region, the determinant of the $\nu + (n - \nu) = n$ linear forms (8_2), (9) in the n impulses p_l is distinct from zero. Thus, from (8_2) and (9),

$$(10) \qquad p_l = \sum_{\lambda=1}^{\nu} \pi_\lambda \Gamma_{l\lambda}(\kappa_1, \cdots, \kappa_\nu); \qquad (l = 1, 2, \cdots, n),$$

where the $n\nu$ functions $\Gamma_{l\mu}$ are uniquely determined and have continuous partial derivatives of the second order. On substituting (8_1) and (10) into (2), one obtains a function H of the κ_λ and π_λ, so that (6) holds in virtue of (8_1) and (10). Now the function H thus defined has the following property: For those solutions of (1) along which (9) is satisfied the canonical system (1) is equivalent to the canonical system (7). The proof requires, as in the case (i), but straightforward differentiations and substitutions,[5] as seen by differentiating (8_1) and (10) with respect to t and (6) with respect to κ_ν, π_ν.

(iii) Let $i = n - m + 1$, $n - m + 2, \cdots, n$, where m is a positive integer less than the degree of freedom of a canonical system (1), and let c_i be a constant and h_i a function of $q_1, \cdots, q_{n-m}$ and $p_1, \cdots, p_{n-m}$ which has continuous partial derivatives of the second order in the region under consideration. Suppose that the m pairs of relations

$$(11) \quad q_i = h_i(q_1, \cdots, q_{n-m}, p_1, \cdots, p_{n-m}), \quad p_i = c_i; \quad (i = n - m + 1, \cdots, n)$$

[5] Cf. also T. Levi-Civita, "Solla introduzione di vincoli olonomi nelle equazioni dinamiche di Hamilton," *Atti del Reale Istituto Veneto*, ser. 8, vol. 18_1 (1916), pp. 387-395.

form an invariant system [6] of (1). On substituting (11) into (2), one obtains a function K of $q_1, \cdots, q_m, p_1, \cdots, p_m$ and the c_i, so that, on considering the numbers c_i as fixed,

$$H = K, \text{ where } K = K(q_1, \cdots, q_m, p_1, \cdots, p_m). \tag{12}$$

Now the function K thus defined has the following property: For those solutions of (1) along which (11) is satisfied the canonical system (1) is equivalent to the canonical system

$$q'_k = \partial K/\partial p_k, \qquad p'_k = -\partial K/\partial q_k; \qquad (k = 1, 2, \cdots, n-m). \tag{13}$$

This is easily verified by noting that

$$\partial H/\partial q_i = 0; \qquad (i = n-m+1, \cdots, n)$$

in virtue of the invariant system (11) of (1).

All three rules (i), (ii), (iii) are to the effect that under the conditions specified the new Hamiltonian function, H or K, is obtained from the original Hamiltonian function H by actual substitution of the transformation formulae. In such a case one says that the transformation formulae define a *completely canonical* transformation. It may be mentioned that all three rules (i), (ii), (iii) are very special cases of a general theorem dealing with canonical systems which are subjected either to constraints or to invariant relations.[7]

A notation. From now on i, j, k will denote subscripts each of which has one of the three values 1, 2, 3. The summation symbol $\Sigma\, a_{ijk}$ will denote the sum

$$\Sigma\, a_{ijk} = a_{123} + a_{231} + a_{312} \qquad (\neq \Sigma\, a_{ikj}),$$

so that, in particular

$$\Sigma\, a_{ij} = a_{12} + a_{23} + a_{31}, \qquad \Sigma\, a_i = a_1 + a_2 + a_3.$$

---

[6] A system of s relations

$$r_h(x_1, \cdots, x_k, t) = 0; \qquad (h = 1, \cdots, s) \tag{*}$$

is an invariant system of a system of k differential equations

$$dx_m/dt = X_m(x_1, \cdots, x_k, t); \qquad (m = 1, \cdots, k), \tag{**}$$

if the functions r_h have continuous partial derivatives of the first order and

$$dr_h(x_1, \cdots, x_k, t)/dt = 0; \qquad (h = 1, \cdots, s)$$

is a consequence of (*) and (**).

[7] Cf. a forthcoming paper where holonomic and certain non-holonomic constraints will be treated by a method used in the case of free canonical systems by E. R. van Kampen and A. Wintner, "On the canonical transformations of Hamiltonian systems," *American Journal of Mathematics*," vol. 58 (1936), pp. 851-863.

An equation of the type $a_{ijk} = b_{ijk}$ will be meant to hold not necessarily for all 6, merely for the 3 cyclic, permutations $(1,2,3)$, $(2,3,1)$, $(3,1,2)$ of $(i,j,k) = (1,2,3)$.

II. *A canonical transformation of a 16-dimensional phase-space.* Let P_0 be an 18-dimensional phase-space of 9 coördinates x_i, y_i, z_i and of 9 impulses X_i, Y_i, Z_i, where $i = 1, 2, 3$. Consider that subspace of P_0 on which

$$(14_1)\qquad \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix} = 0; \qquad (14_2)\qquad \Sigma \frac{1}{m_i} \begin{vmatrix} X_i & Y_i & Z_i \\ x_j & y_j & z_j \\ x_k & y_k & z_k \end{vmatrix} = 0,$$

where m_1, m_2, m_3 are three positive constants. Suppose further that

$$(15_1)\qquad r = 2; \qquad (15_2)\qquad \Sigma z_i^2 \neq 0,$$

where r denotes the rank of the matrix of (14_1). The subspace of P_0 characterized by the four conditions (14_1), (14_2), (15_1), (15_2) will be denoted by P, so that P is 16-dimensional.

Let Π denote the 16-dimensional phase-space of 8 coördinates $\nu, \iota; \xi_i, \eta_i$ such that

$$(16_1)\qquad \Delta \neq 0; \qquad (16_2)\qquad \sin \iota \neq 0$$

and of 8 impulses $\mathrm{N}, \mathrm{I}; \Xi_i, \mathrm{H}_i$, where $i = 1, 2, 3$ and Δ is an abbreviation for

$$(17)\qquad \Delta = \Sigma\, m_i m_j (\xi_i \eta_j - \xi_j \eta_i)^2.$$

One can express condition (16_1) without reference to the three constants $m_i > 0$ by requiring that the two vectors (ξ_1, ξ_2, ξ_3), (η_1, η_2, η_3) be linearly independent. It is clear from (17) that

$$(18)\qquad \Delta = \mathrm{M}_{\xi\xi}\mathrm{M}_{\eta\eta} - \mathrm{M}_{\xi\eta}^2,$$

where

$$(19)\qquad \mathrm{M}_{\xi\xi} = \Sigma\, m_i \xi_i^2, \qquad \mathrm{M}_{\xi\eta} = \Sigma\, m_i \xi_i \eta_i, \qquad \mathrm{M}_{\eta\eta} = \Sigma\, m_i \eta_i^2.$$

Now put

$$(20_1)\qquad \begin{aligned} x_i &= \xi_i \cos \nu - \eta_i \sin \nu \cos \iota, \\ y_i &= \xi_i \sin \nu + \eta_i \cos \nu \cos \iota, \\ z_i &= \xi_i \cdot 0 \quad + \eta_i \sin \iota, \end{aligned}$$

and

$$(20_2)\qquad \begin{aligned} \mathrm{N} &= \Sigma\, \xi_i (-X_i \sin \nu + Y_i \cos \nu) - \cos \iota\, \Sigma\, \eta_i (X_i \cos \nu + Y_i \sin \nu), \\ \mathrm{I} &= \Sigma\, \eta_i (X_i \sin \nu \sin \iota - Y_i \cos \nu \sin \iota + Z_i \cos \iota), \end{aligned}$$

finally

$$(20_3)\qquad \begin{aligned} \Xi_i &= X_i \cos\nu + Y_i \sin\nu, \\ \mathrm{H}_i &= -X_i \sin\nu\cos\iota + Y_i \cos\nu\cos\iota + Z_i \sin\iota, \end{aligned}$$

where $i = 1, 2, 3$. It will be shown that the formulae (20_1), (20_2), (20_3) define a completely canonical transformation of the two 16-dimensional phase-spaces P, Π into each other.

First, if G_x, G_y, G_z and Γ are defined by

$$(21)\qquad G_x = \Sigma(y_i Z_i - z_i Y_i),\ G_y = \Sigma(z_i X_i - x_i Z_i),\ G_z = \Sigma(x_i Y_i - y_i X_i)$$

and

$$(22)\qquad \Gamma = \Sigma(\xi_i \mathrm{H}_i - \eta_i \Xi_i),$$

then

$$(23_1)\qquad \mathrm{N} = G_z, \qquad \mathrm{I} = G_x \cos\nu + G_y \sin\nu,$$

$$(23_2)\qquad \Gamma = G_x \sin\nu\sin\iota - G_y \cos\nu\sin\iota + G_z \cos\iota$$

are identities in virtue of (20_1), (20_2), (20_3). This is seen by expressing G_x, G_y, G_z and Γ by means of (20_1), (21) and (20_3), (22) in terms of ν, ι; ξ_i, η_i; X_i, Y_i, Z_i and substituting the result and (20_2) into (23_1), (23_2).

Next, on placing

$$(24)\qquad \mathrm{T}_i = X_i \sin\nu\sin\iota - Y_i \cos\nu\sin\iota + Z_i \cos\iota,$$

where $i = 1, 2, 3$, it is clear that (20_3) and (24) determine, for every fixed i, an orthogonal transformation of the vector (X_i, Y_i, Z_i) into the vector $(\Xi_i, \mathrm{H}_i, \mathrm{T}_i)$, so that, the inverse transformation being obtained by transposing the matrix,

$$(25)\qquad \begin{aligned} X_i &= \Xi_i \cos\nu - \mathrm{H}_i \sin\nu\cos\iota + \mathrm{T}_i \sin\nu\sin\iota, \\ Y_i &= \Xi_i \sin\nu + \mathrm{H}_i \cos\nu\cos\iota - \mathrm{T}_i \cos\nu\sin\iota, \\ Z_i &= \Xi_i \cdot 0 \quad + \mathrm{H}_i \sin\iota \qquad + \mathrm{T}_i \cos\iota. \end{aligned}$$

Furthermore,

$$(26_1)\qquad \Sigma\,\eta_i \mathrm{T}_i = \mathrm{I}; \qquad \Sigma\,\xi_i \mathrm{T}_i = \frac{\Gamma\cos\iota - \mathrm{N}}{\sin\iota}, \text{ if } \sin\iota \neq 0$$

are identities, while

$$(26_2)\qquad \Sigma\, m_j m_k (\xi_j \eta_k - \xi_k \eta_j)\mathrm{T}_i = 0$$

is an identity in virtue of (14_2). In fact, the first of the relations (26_1) is obvious from the second of the relations (20_2) and from (24), while the second of the relations (26_1) follows by substituting (20_3) into (22) and using (20_1) and (24). Finally, on substituting (20_1) and (25) into the determinants occurring in (14_2), a straightforward calculation shows that (14_2) may be written in the form (26_2).

Now it is easily verified by direct substitution that (14_1), (15_1) and (15_2) are implied by (20_1), (16_1) and (16_2). Conversely, if 9 given values x_i, y_i, z_i, where $i = 1, 2, 3$, satisfy (14_1), (15_1) and (15_2), then the equations (20_1), (16_1) and (16_2) may be satisfied by 8 values ν, ι, ξ_i, η_i and, on disregarding the trigonometrical ambiguity of the angles ν and ι, the 8 values ν, ι, ξ_i, η_i thus obtained are uniquely determined by the 9 values x_i, y_i, z_i. This is seen from the fact that the coefficients of the linear substitution (20_1) form the first and second columns of a three-rowed orthogonal matrix.

Accordingly, conditions (14_1), (15_1) and (15_2) are equivalent to conditions (16_1) and (16_2) in virtue of (20_1) alone. On the other hand, on identifying (20_1) with (8_1), it is seen by partial differentiation of (20_1) that the system (20_2), (20_3) goes over into (8_2). Hence, in order to conclude from the rule (ii) of Section I that (20_1), (20_2), (20_3) determine a completely canonical transformation of the two 16-dimensional phase-spaces P, Π into each other, it is sufficient to prove that the 9 linear relations (14_2), (20_2), (20_3) for the 9 variables X_i, Y_i, Z_i have a non-vanishing determinant. Now it is seen from (24) and (25) that this 9-rowed determinant cannot vanish, unless so does the 3-rowed determinant of the 3 linear conditions (26_1), (26_2) for the 3 variables T_i. Finally, this 3-rowed determinant is easily verified to be identical with the expression (17) and is, therefore, distinct from zero in view of the assumption (16_1).

Let it be noted for later application that

$$(27) \qquad \Sigma\, m_i^{-1}(X_i^2 + Y_i^2 + Z_i^2) = \Sigma\, m_i^{-1}(\Xi_i^2 + \mathrm{H}_i^2) + \Sigma\, m_i^{-1}\mathrm{T}_i^2$$

and

$$(28) \qquad \Sigma\, m_i^{-1}\mathrm{T}_i^2\Delta = \mathrm{I}^2\mathrm{M}_{\xi\xi} - 2\mathrm{I}\frac{\Gamma\cos\iota - \mathrm{N}}{\sin\iota}\mathrm{M}_{\xi\eta} + \frac{(\Gamma\cos\iota - \mathrm{N})^2}{\sin^2\iota}\mathrm{M}_{\eta\eta}.$$

In fact, since (25) is an orthogonal substitution, (27) is obvious. On the other hand, on solving the 3 linear equations (26_1), (26_2) with respect to T_1, T_2, T_3 and using the definitions (17) and (19), one obtains

$$m_i^{-1}\mathrm{T}_i\Delta = \mathrm{I}(\eta_i\mathrm{M}_{\xi\xi} - \xi_i\mathrm{M}_{\xi\eta}) - \frac{\Gamma\cos\iota - \mathrm{N}}{\sin\iota}(\eta_i\mathrm{M}_{\xi\eta} - \xi_i\mathrm{M}_{\eta\eta}).$$

Hence (28) follows by direct substitution, if one uses (18) and (19).

III. *A canonical transformation of an 8-dimensional phase-space.* Let Ψ_0 denote the 12-dimensional phase-space of 6 coördinates ξ_i, η_i and of 6 impulses Ξ_i, H_i, where $i = 1, 2, 3$. For 3 given positive constants m_i, let Ψ denote that subspace of Ψ_0 on which

$$(29_1) \qquad \Sigma\, m_i\xi_i = 0, \qquad \Sigma\, m_i\eta_i = 0;$$

$$(29_2) \qquad \Sigma\, \Xi_i = 0, \qquad \Sigma\, \mathrm{H}_i = 0;$$

$$(29_3) \qquad \delta > 0, \text{ where } \delta = \tfrac{1}{2}\begin{vmatrix} \xi_1 & \eta_1 & 1 \\ \xi_2 & \eta_2 & 1 \\ \xi_3 & \eta_3 & 1 \end{vmatrix},$$

so that the space Ψ is 8-dimensional. Condition (29_3) may be expressed by saying that the 3 points (ξ_i, η_i) form in a (ξ, η)-plane a positively oriented triangle of area $\delta \neq 0$. Let ϕ_i denote the inclination of that oriented side of this triangle which is opposite to the vertex (ξ_i, η_i), so that

$$(30) \qquad \cos\phi_i = \frac{\xi_k - \xi_j}{[(\xi_k - \xi_j)^2 + (\eta_k - \eta_j)^2]^{\frac{1}{2}}}, \quad \sin\phi_i = \frac{\eta_k - \eta_j}{[(\xi_k - \xi_j)^2 + (\eta_k - \eta_j)^2]^{\frac{1}{2}}},$$

where $[\ ]^{\frac{1}{2}} > 0$. Furthermore,

$$(31) \qquad \theta_i = \phi_k - \phi_j,$$

where θ_i denotes the oriented exterior angle at the vertex (ξ_i, η_i) of the triangle, so that

$$(32) \qquad \Sigma\, \theta_i = 0$$

and not only $\Sigma\, \theta_i \equiv 0 \pmod{2\pi}$. The notations (17), (19) of Section II will be used in the space Ψ also. On multiplying the i-th row of the determinant (29_3) by $m_i^{\frac{1}{2}}$ and calculating the square of the resulting determinant by column-by-column multiplication, it is easily seen from (18), (19) and (29_1) that

$$(33) \qquad \Delta = 4\mu\delta^2,$$

where μ is the positive constant

$$(34) \qquad \mu = m_1 m_2 m_3/(m_1 + m_2 + m_3).$$

Let Θ denote the 8-dimensional phase-space of 4 coördinates ω, ρ_i such that

$$(35) \qquad 0 < \rho_i < \rho_j + \rho_k$$

and of 4 impulses Ω, P_i, where $i = 1, 2, 3$. Condition (35) means that the ρ_i are the sides of a non-degenerate triangle.

Now it will be shown that the formulae

$$(36_1) \qquad \omega = \tfrac{1}{3}\Sigma\, \phi_i,$$

$$(36_2) \qquad \rho_i = [(\xi_k - \xi_j)^2 + (\eta_k - \eta_j)^2]^{\frac{1}{2}},$$

$$(36_3)\qquad \begin{aligned} \Xi_i &= \mathrm{P}_j \cos\phi_j - \mathrm{P}_k \cos\phi_k - \frac{\Omega}{3}\left(\frac{\sin\phi_j}{\rho_j} - \frac{\sin\phi_k}{\rho_k}\right), \\ \mathrm{H}_i &= \mathrm{P}_j \sin\phi_j - \mathrm{P}_k \sin\phi_k + \frac{\Omega}{3}\left(\frac{\cos\phi_j}{\rho_j} - \frac{\cos\phi_k}{\rho_k}\right), \end{aligned}$$

define in virtue of (30) a completely canonical transformation of the two 8-dimensional phase-spaces Ψ, Θ into each other.

First, (36_2) means the identification of the two triangles mentioned above. Hence

$$(37)\qquad \cos\theta_i = \frac{\rho_i^2 - \rho_j^2 - \rho_k^2}{2\rho_j\rho_k}, \qquad \sin\theta_i = \frac{2\delta}{\rho_j\rho_k},$$

where δ is the area (29_3), so that

$$(38)\qquad 16\delta^2 = (\rho_1 + \rho_2 + \rho_3)(-\rho_1 + \rho_2 + \rho_3)(\rho_1 - \rho_2 + \rho_3)(\rho_1 + \rho_2 - \rho)\,;\ \delta > 0.$$

Furthermore, from (36_1) and (31),

$$(39)\qquad \phi_i = \omega + \tfrac{1}{3}(\theta_j - \theta_k),$$

while from (36_2) and (30)

$$(40)\qquad \xi_k - \xi_j = \rho_i \cos\phi_i, \qquad \eta_k - \eta_j = \rho_i \sin\phi_i.$$

It is seen from (40) and (29_1) that

$$(41)\qquad \begin{aligned} (m_1 + m_2 + m_3)\xi_i &= m_k\rho_j \cos\phi_j - m_j\rho_k \cos\phi_k \\ (m_1 + m_2 + m_3)\eta_i &= m_k\rho_j \sin\phi_j - m_j\rho_k \sin\phi_k. \end{aligned}$$

Now (39) represents in view of (37) and (38) the 3 angles ϕ_i in terms of the 4 coördinates ω, ρ_i. Thus (41) represents the 6 coördinates ξ_i, η_i of the phase-space Ψ in terms of the 4 coördinates ω, ρ_i of the phase-space Θ. Accordingly, on identifying the system (36_1), (36_2) with (3_1) and (29_1) with (4), it is clear that the Jacobian mentioned under (i) in Section I is distinct from zero. Finally, partial differentiation of (36_1), (36_2) shows that (36_3) goes over into (3_2) in view of (30). This proves that (36_1), (36_2), (36_3) define a completely canonical transformation. Incidentally, (36_3) clearly implies (29_2).

Let it be mentioned for later application that

$$(42)\qquad \begin{aligned} \Sigma\, m_i^{-1}(\Xi_i^2 + \mathrm{H}_i^2) &= \Sigma\, m_i^{-1}(\mathrm{P}_j^2 + \mathrm{P}_k^2 - 2\mathrm{P}_j\mathrm{P}_k \cos\theta_i) \\ &\quad + \frac{2\Omega}{3}\Sigma\, m_i^{-1}\left(\frac{\mathrm{P}_j}{\rho_k} - \frac{\mathrm{P}_k}{\rho_j}\right)\sin\theta_i \\ &\quad + \frac{\Omega^2}{9}\Sigma\, m_i^{-1}\left(\frac{1}{\rho_j^2} + \frac{1}{\rho_k^2} - \frac{2\cos\theta_i}{\rho_j\rho_k}\right). \end{aligned}$$

In fact, on calculating $\Xi_i^2 + H_i^2$ from (36_3) and using then (31), the identity (42) follows. Similarly,

$$\Gamma = \Omega, \tag{43}$$

where Γ is defined by (22). In fact, on substituting (36_3) into (22) and using then (40), one clearly obtains (43). Finally,

$$\mathbf{M}_{\eta\eta} = \mu \Sigma\, m_i^{-1}\rho_i^2 \sin^2 [\omega + \tfrac{1}{3}(\theta_k - \theta_j)]. \tag{44}$$

In fact, since

$$\Sigma\, m_i\eta_i^2 = \mu \Sigma\, m_i^{-1}(\eta_k - \eta_j)^2$$

is an identity in virtue of (29_1) and (34), it is seen from (39) and (40) that (44) is equivalent to the definition (19) of $\mathbf{M}_{\eta\eta}$. Similar formulae hold for $\mathbf{M}_{\xi\xi}$ and $\mathbf{M}_{\xi\eta}$.

IV. *Application to the non-planar problem of three bodies.* The completely canonical substitutions given in Sections II and III may be applied to a symmetrical reduction of the degree of freedom of the non-planar problem of three bodies from 9 to 4, using merely the conservation of the center of mass and of the angular momentum.

Let x_i, y_i, z_i be the Cartesian coördinates of the mass m_i in a barycentric inertial coördinate system (x, y, z), and let the plane $z = 0$ be chosen such that its positive normal has the direction of the vector which represents the constant angular momentum, this vector being not the zero vector in view of the assumption that the motion is non-planar. If C denotes the length of this vector and (X_i, Y_i, Z_i) the impulse vector $(m_i x'_i, m_i y'_i, m_i z'_i)$ of m_i, the conservation of the angular momentum is represented by the invariant relations

$$G_x = 0, \quad G_y = 0, \quad G_z = C, \tag{45}$$

where G_x, G_y, G_z are defined by (21). The preservation of the center of mass is expressed by the invariant relations

$$\Sigma\, m_i x_i = 0, \qquad \Sigma\, m_i y_i = 0, \qquad \Sigma\, m_i z_i = 0; \tag{46$_1$}$$

$$\Sigma\, X_i = 0, \qquad \Sigma\, Y_i = 0, \qquad \Sigma\, Z_i = 0, \tag{46$_2$}$$

since (x, y, z) is a barycentric inertial coördinate system. Finally, the Hamiltonian function is

$$H = \tfrac{1}{2}\Sigma\, m_i^{-1}(X_i^2 + Y_i^2 + Z_i^2) - \Sigma\, m_j m_k \rho_i^{-1}, \tag{47}$$

where

$$\rho_i = [(x_k - x_j)^2 + (y_k - y_j)^2 + (z_k - z_j)]^{\frac{1}{2}}. \tag{48}$$

Since it is assumed that the motion is non-planar, and since the equations of motion define the coördinates and impulses as analytic functions of the time t, it is clear [8] that those values of t for which the three masses are either collinear or lie in the invariable plane $z = 0$ have no cluster point on the t-axis and can, therefore, be excluded from the considerations of this Section IV. Thus conditions (15_1), (15_2) are satisfied. Furthermore, (14_1) and (14_2) are implied by (46_1) and (46_2), since $m_i > 0$.

Now introduce in place of the barycentric inertial coördinate system (x, y, z) a barycentric but non-inertial coördinate system (ξ, η, ζ) in such a way that the plane $\zeta = 0$ is, for every t, the plane of the triangle formed by the three masses and that the ξ-axis is in the invariable plane $z = 0$ for every t. Then, if ν denotes the node and ι the inclination of the moving plane $\zeta = 0$ with reference to the fixed plane $z = 0$, the Eulerian representation of a rotation shows that the transformation formulae of the coördinates are given by (20_1). Furthermore, (16_1) and (16_2) are satisfied, since the three masses are neither collinear nor all contained in the (x, y)-plane. Now (20_1), (20_2), (20_3) were shown (Section II) to define a completely canonical transformation. Hence, on expressing the Hamiltonian function (47) by means of the transformation formulae in terms of the coördinates ν, ι, ξ_i, η_i and the canonically conjugate variables N, I, Ξ_i, H_i, and denoting the resulting function again by H, the equations of motion are

$$(49_1)\qquad \nu' = \frac{\partial H}{\partial \mathrm{N}}, \quad \mathrm{N}' = -\frac{\partial H}{\partial \nu}; \quad \iota' = \frac{\partial H}{\partial \mathrm{I}}, \quad \mathrm{I}' = -\frac{\partial H}{\partial \iota},$$

$$(49_2)\qquad \xi'_i = \frac{\partial H}{\partial \Xi_i}, \quad \Xi'_i = -\frac{\partial H}{\partial \xi_i}; \quad \eta'_i = \frac{\partial H}{\partial \mathrm{H}_i}, \quad \mathrm{H}'_i = -\frac{\partial H}{\partial \eta_i},$$

where

$$(50)\qquad H = \tfrac{1}{2}\Sigma\, m_i^{-1}(\Xi_i^2 + \mathrm{H}_i^2) - \Sigma\, m_j m_k \rho_i^{-1} + \frac{1}{2\Delta}\left[\mathrm{I}^2 \mathrm{M}_{\xi\xi} - 2\mathrm{I}\,\frac{\Gamma\cos\iota - \mathrm{N}}{\sin\iota}\,\mathrm{M}_{\xi\eta} + \frac{(\Gamma\cos\iota - \mathrm{N})^2}{\sin^2\iota}\,\mathrm{M}_{\eta\eta}\right]$$

in view of (47), (27), (28). It is understood that in (50)

$$\Gamma,\ \Delta,\ \mathrm{M}_{\xi\xi},\ \mathrm{M}_{\xi\eta},\ \mathrm{M}_{\eta\eta} \text{ and } \rho_i$$

are thought of as expressed by means of (22), (17), (19) and (36_2) in terms of ξ_i, η_i, Ξ_i, H_i.

In order to reduce the degree of freedom of the simultaneous system (49_1), (49_2) from 8 to 6 by means of the invariant relations (45), one can separate

[8] In fact, it is well known that if a solution is collinear, then it is necessarily planar. Cf.[12]

the 4 equations (49_1) from the 12 equations (49_2) as follows: First, from (23_1), (23_2) and (45),

$$(51) \qquad \mathbf{N} = C, \qquad \mathbf{I} = 0, \qquad \cos \iota = \Gamma/C,$$

while the fourth of the variables belonging to (49_1), namely the node ν, does not occur in the Hamiltonian function (50). This fact and (51) represent an invariant system of $2m = 4$ relations of the type (11). Thus it is seen from the rule (iii) of Section I that (49_2) may be replaced by

$$(52) \qquad \xi'_i = \frac{\partial K}{\partial \Xi_i}, \quad \Xi'_i = -\frac{\partial K}{\partial \xi_i}; \qquad \eta'_i = \frac{\partial K}{\partial \mathrm{H}_i}, \quad \mathrm{H}'_i = -\frac{\partial K}{\partial \eta_i},$$

where the Hamiltonian function

$$(53_1) \qquad K = K(\xi_1, \cdots, \eta_3, \Xi_1, \cdots, \mathrm{H}_3)$$

is obtained from (50) by substituting (51) into (50). Consequently,[9]

$$(53_2) \qquad K = \tfrac{1}{2}\Sigma\, m_i^{-1}(\Xi_i^2 + \mathrm{H}_i^2) - \Sigma\, m_j m_k \rho_i^{-1} + \tfrac{1}{2}(C^2 - \Gamma^2)\mathrm{M}_{\eta\eta}/\Delta,$$

where Γ, Δ, $\mathrm{M}_{\xi\eta}$ and ρ_i are thought of as expressed by means of (22), (17), (19) and (36_2) in terms of ξ_i, η_i, Ξ_i, H_i. The conservative canonical system (52) with 6 degrees of freedom possesses the 4 invariant relations (29_1), (29_2), as easily verified either from (20_1), (20_2) and (46_1), (46_2) or by partial differentiation of (53_2).

Now the degree of freedom of the system (52) may be reduced by means of the invariant relations (29_1), (29_2) from 6 to 4 by using the canonical transformation treated in Section III. In fact, (29_3) is satisfied in view of the assumption that the three masses m_i are not colinear. Hence the result of Section III is applicable and shows that (52) may be replaced by the 4 pairs of simultaneous equations

$$(54_1) \qquad \omega' = \frac{\partial \mathrm{K}}{\partial \Omega}, \qquad \Omega' = -\frac{\partial \mathrm{K}}{\partial \omega},$$

$$(54_2) \qquad \rho'_i = \frac{\partial \mathrm{K}}{\partial \mathrm{P}_i}, \qquad \mathrm{P}'_i = -\frac{\partial \mathrm{K}}{\partial \rho_i}, \qquad (i = 1, 2, 3),$$

where the Hamiltonian function

$$(55_1) \qquad \mathrm{K} = \mathrm{K}(\omega, \rho_1, \rho_2, \rho_3, \Omega, \mathrm{P}_1, \mathrm{P}_2, \mathrm{P}_3)$$

[9] This agrees with a result of Levi-Civita, "Sulla riduzione del problema dei tre corpi," *Atti del Reale Istituto Veneto*, ser. 8, vol. 17_1 (1915), pp. 907-939, more particularly p. 923, formula (28).

is obtained by expressing the Hamiltonian function (53_1) by means of (36_1), (36_2), (36_3) and (30) in terms of ω, ρ_i, Ω, P_i, so that

$$(55_2)\quad \mathrm{K} = \Sigma \frac{1}{2m_i}(\mathrm{P}_j^2 + \mathrm{P}_k^2 - 2\mathrm{P}_j\mathrm{P}_k\cos\theta_i) + \frac{\Omega}{3}\Sigma\left(\frac{\mathrm{P}_j}{\rho_k} - \frac{\mathrm{P}_k}{\rho_j}\right)\frac{\sin\theta_i}{m_i} + \frac{\Omega^2}{9}\Sigma\frac{1}{m_i}\left(\frac{1}{\rho_j^2} + \frac{1}{\rho_k^2} - \frac{2\cos\theta_i}{\rho_j\rho_k}\right) + \frac{C^2-\Omega^2}{4\delta^2}\Sigma\frac{\rho_i^2}{m_i}\sin^2[\omega + \tfrac{1}{3}(\theta_j - \theta_k)] - \Sigma\frac{m_j m_k}{\rho_i}$$

in view of (42), (43), (44), (33), (34) and (53_2). It is understood that in (55_2) one has to express $\theta_1, \theta_2, \theta_3$ and δ by means of (37), (32) and (38) in terms of the three ρ_i. Thus (54_1), (54_2) is a conservative canonical system with 4 degrees of freedom, and with a Hamiltonian function (55_2) which is entirely symmetrical with respect to the three masses.

Remark. If a solution of the system (54_1), (54_2) is known, the corresponding solution of (52) follows from (36_1), (36_2), (36_3) and (30), while the functions

$$\nu = \nu(t), \qquad \mathrm{N} = \mathrm{N}(t), \qquad \iota = \iota(t), \qquad \mathrm{I} = \mathrm{I}(t)$$

follow by a single quadrature. In fact, if ξ_i, η_i, Ξ_i, H_i are known functions of t, then so are $\mathrm{N}\ (\equiv C)$, $\mathrm{I}\ (\equiv 0)$ and ι in view of (51) and (22). Furthermore, on substituting (50) into the first of the equations (49_1) and using then (51), one obtains

$$(56)\qquad \nu' = C\mathrm{M}_{\eta\eta}/\Delta,$$

where $\mathrm{M}_{\eta\eta}$ and Δ are, in view of (17) and (19), known functions of t. Hence $\nu = \nu(t)$ follows from (56) by a quadrature.[10] This corresponds to the fact that the node ν is, in view of (50), an ignorable coördinate in (49_1), (49_2), a fact usually referred to as "the elimination of the node" (Lagrange-Jacobi). The relation which corresponds to (56) in case of the inclination is

$$(57)\qquad \iota'/\sin\iota = C\mathrm{M}_{\xi\eta}/\Delta,$$

a relation obtained in the same way as (56).

V. *Application to the planar problem of three bodies.* It has been assumed so far that the solution is non-planar. Now let the solution be planar

[10] Cf. T. Levi-Civita, *loc. cit.*, [9], pp. 931-933; also Maria Ronchi, "Sulla riduzione esplicita del problema dei tre corpi," *Atti Reale Istituto Veneto*, ser. 9, vol. 1_2 (1917), pp. 1221-1225.

and let (ξ, η) be a barycentric inertial plane containing all three paths. Then, on excluding the case of collinear configurations, (29_1), (29_2), (29_3) are satisfied. Hence the Hamiltonian function is $H = T - U$, where $2T$ is represented by (42), while $U = \Sigma\, m_j m_k \rho_i^{-1}$. It is seen from (22) and (43) that the preservation of the angular momentum is the relation

$$\Omega = C. \tag{58}$$

On substituting (58) into (42) and applying the rule (iii) of Section I to the canonically conjugate pair Ω, ω, where now the coördinate ω does not occur explicitly, it follows that the equations of motion may be written in the form (54_2), where K is given not by (55_2) but by the expression [11]

$$\mathrm{K} = \tfrac{1}{2}\Sigma\, m_i^{-1}(\mathrm{P}_j^2 + \mathrm{P}_k^2 - 2\mathrm{P}_j\mathrm{P}_k \cos\theta_i) - \Sigma\, m_j m_k/\rho_i + \frac{C}{3}\Sigma\left(\frac{\mathrm{P}_j}{\rho_k} - \frac{\mathrm{P}_k}{\rho_j}\right)\frac{\sin\theta_i}{m_i} + \frac{C^2}{9}\Sigma\,\frac{1}{m_i}\left(\frac{1}{\rho_j^2} + \frac{1}{\rho_k^2} - \frac{2\cos\theta_i}{\rho_j\rho_k}\right). \tag{59}$$

If a solution of this conservative problem (54_2), which has three degrees of freedom, is known, the angle $\omega = \omega(t)$ defined by (36_1) and (30) follows by a quadrature, the reason being that ω is an ignorable coördinate in view of (58).

Remark. Collinear configurations of the three masses have been excluded so far. Such positions (syzygies) occur only for isolated values of t, unless the configuration is collinear for every t. In the latter case the solution is necessarily planar, since [12] the solution is then a homothetic solution, unless the line containing the three masses is independent of t.

THE JOHNS HOPKINS UNIVERSITY.

[11] This agrees with the result of Murnaghan, *loc. cit.*, [3], p. 832, since θ_i in (59) and Murnaghan's A_i are exterior and interior angles respectively.

[12] Cf., e. g., Ch. H. Müntz, "Die Aehnlichkeitsbewegungen beim allgemeinen n-Körperproblem," *Mathematische Zeitschrift*, vol. 15 (1923), pp. 169-187.

ON THE VALUES OF THE RIEMANN ζ-FUNCTION ON FIXED LINES $\sigma > 1$.[1]

By RICHARD KERSHNER.

For a fixed $\sigma > 1$, let $M(\sigma)$ denote the closure of the set of values taken by the almost periodic function

$$-\log \zeta(s) = \sum_{n=1}^{\infty} \log\,(1 - p_n^{-\sigma} e^{-it \log p_n}), \qquad -\infty < t < +\infty,$$

where $\zeta(s) = \zeta(\sigma + it)$ is the Riemann ζ-function. It has been shown by Bohr [1] that, due to the linear independence of the logarithms of the prime numbers, $M(\sigma)$ is the vectorial sum [2] of the sequence of curves $S(p_1^{-\sigma})$, $S(p_2^{-\sigma}), \cdots$, where $S(r)$ denotes, for a fixed $r < 1$, the curve

$$w = u + iv = \log\,(1 + re^{i\theta}), \qquad 0 \leqq \theta < 2\pi, \tag{1}$$

in the w-plane. Since every $S(r)$ is easily seen to be a convex Jordan curve, it follows from Bohr's general results (cf. [2]) on the vectorial addition of convex curves that $M(\sigma)$ is,[3] for every fixed $\sigma > 1$, either a closed bounded convex region bordered by a convex Jordan curve $A(\sigma)$ or a closed ring-shaped region bordered by two convex Jordan curves $A(\sigma)$ and $B(\sigma)$, where $B(\sigma)$ is entirely within $A(\sigma)$. Following a suggestion of Wintner, Haviland [5] applied the supporting function (Stützfunktion) of Brunn and Minkowski to the study of $A(\sigma)$, using the fact that the supporting function of $A(\sigma)$ is the sum of the supporting functions of all $S(p_n^{-\sigma})$. A corresponding study of the inner curve $B(\sigma)$, in case $B(\sigma)$ exists, has recently been made by the author [8] and results similar to, but more complicated than, those concerning $A(\sigma)$ were obtained. Actually these investigations treated not only the particular problem of $M(\sigma)$ but the general problem of the vectorial addition of any sequence of convex curves. The object of the present paper is to apply

[1] Received October 28, 1936.

[2] By the vectorial sum $S_1(+)\,S_2(+)\cdots(+)\,S_n$ of n sets S_j of complex numbers is meant the set of numbers w representable in at least one way in the form $w = w_1 + w_2 + \cdots + w_n$ where $w_j \subset S_j$. If $n = \infty$, one means by $S_1(+)S_2(+)\cdots$ the set of points w for which $w = w_1 + w_2 + \cdots$ is a convergent series and $w_j \subset S_j$. In the case under consideration ($\sigma > 1$) this series is absolutely convergent for every choice of $w_j \subset S_j$.

[3] For a short proof of this statement cf. Jessen and Wintner [6], p. 69.

these general results to a detailed discussion of $M(\sigma)$ for varying values of σ. Since the closure $N(\sigma)$ of the values attained by $\zeta(\sigma+it)$ on a fixed line $\sigma>1$ may be obtained from $M(\sigma)$ by a trivial exponential mapping, the results to be obtained may easily be formulated in terms of $\zeta(s)$ instead of $\log\zeta(s)$.

Mention should first be made of the following result of Bohr and Jessen [4]:

(I) *There exists a constant* $\sigma_0>1$ *such that* $M(\sigma)$ *is or is not a convex region* (i.e., $B(\sigma)$ *does not or does exist*) *according as* $1<\sigma\leqq\sigma_0$ *or* $\sigma_0<\sigma<+\infty$. *This* σ_0 *is the only root* $\sigma>1$ *of the equation*

$$\text{(2a)}\qquad \arcsin 2^{-\sigma}=\sum_{n=2}^{\infty}\arcsin p_n^{-\sigma},\ \textit{so that}\ \sigma_0=1.764\cdots.$$

It is easy to see that $S(r)$ has, for every fixed $r<1$, the two lines of symmetry $v=0$ and $u=\frac{1}{2}\log(1-r^2)$, so that $M(\sigma)$ has, for every fixed $\sigma>1$, the two lines of symmetry

$$v=0 \text{ and } u=u_\sigma=\sum_{n=1}^{\infty}\tfrac{1}{2}\log(1-p_n^{-2\sigma}).$$

It has been shown by Kershner and Wintner [7] that $A(\sigma)$ is, for every $\sigma>1$, a regular analytic curve and that the same is true of $B(\sigma)$ up to its possible corners, provided that $B(\sigma)$ exists. In view of the fact (cf. (II) below) that $B(\sigma)$ has at most two corners for any σ, this statement implies a remark of Bohr and Jessen [4] to the effect that $B(\sigma)$ contains no straight segments for any σ.

The following theorem will now be deduced:

(II) *There exists a constant* $\sigma^0>\sigma_0$ *such that* $B(\sigma)$ *has exactly two corners or no corners according as* $\sigma_0<\sigma<\sigma^0$ *or* $\sigma^0\leqq\sigma<+\infty$. *This* σ^0 *is the only root* $\sigma>1$ *of the equation*

$$\text{(2b)}\qquad 2^{-\sigma}=\sum_{n=2}^{\infty}p_n^{-\sigma},\ \textit{so that}\ \sigma^0=1.778\cdots.$$

If $\sigma_0<\sigma<\sigma^0$, *both corners of* $B(\sigma)$ *lie on the real axis.*

This result has been announced, without proof, by Bohr and Jessen at the end of their paper [4], where it is stated that the proof of (II) is similar to that of (I). In attempting to verify (II), it was found that (II) may be deduced from the results of the paper [8] as follows:

Let $P(r)$ denote, for a fixed r in $0<r<1$, the circle $|z-1|=r$, so

that $S(r)$ is, by (1), the image of $P(r)$ under the conformal mapping $u + iv = \log z$. Thus, if $z = \rho e^{i\phi}$, $-\frac{1}{2}\pi \leqq \phi \leqq \frac{1}{2}\pi$, then,

$$u = \log \rho, \qquad v = \phi. \tag{3}$$

The ranges of ρ and ϕ, when z describes $P(r)$, are

$$1 - r \leqq \rho \leqq 1 + r \quad \text{and} \quad -\arcsin r \leqq \phi \leqq \arcsin r.$$

Let $\alpha(r, v)$, for $0 < r < 1$, $0 < v \leqq \arcsin r$, denote the angle measured in the positive direction from the u-axis to the normal to $S(r)$ at that point of $S(r)$ cut by the line of constant ordinate v for which u is greater, and let $\rho(r, \alpha)$ be the radius of curvature of $S(r)$ at this same point. Similarly, let $\alpha_\sigma(v)$ be the angle measured in the positive direction from the u-axis to the normal to $B(\sigma)$ at that point of $B(\sigma)$ cut by the line of constant ordinate v for which u is greater, and let $\rho_\sigma(\alpha)$ be the radius of curvature of $B(\sigma)$ at this point, if the radius of curvature is defined there. Clearly, $0 < \alpha_\sigma(v) \leqq \frac{1}{2}\pi$, for the section of $B(\sigma)$ for which $\alpha_\sigma(v)$ has been defined. Due to the above-mentioned symmetry of $B(\sigma)$ it will be sufficient to consider, in the sequel, only that arc of $B(\sigma)$ for which $0 < \alpha_\sigma(v) \leqq \frac{1}{2}\pi$.

It has been shown in [8] that if $\rho_\sigma(\alpha)$ is defined for a given α in $0 < \alpha \leqq \frac{1}{2}\pi$, then, since $\rho(r, \alpha) = \rho(r, \alpha + \pi)$,

$$\rho_\sigma(\alpha) = \rho(2^{-\sigma}, \alpha) - \sum_{n=2}^{\infty} \rho(p_n^{-\sigma}, \alpha). \tag{4}$$

It was also shown in [8] that, if $\sigma > 1$ is fixed, the radius of curvature $\rho_\sigma(\alpha)$ exists, and so (4) is valid for all positive $\alpha \leqq \frac{1}{2}\pi$ if and only if the expression on the right of (4) is non-negative for all these α and vanishes for isolated values of α, at most. According to Bohr and Jessen, (cf. [3], p. 402),

$$\rho(r, \alpha) = r/\cos v(r, \alpha),$$

where $v(r, \alpha)$ is the unique inverse in $0 < \alpha \leqq \frac{1}{2}\pi$ of $\alpha(r, v)$, so that $0 < v(r, \alpha) \leqq \arcsin r < \frac{1}{2}\pi$. Finally,

$$\cos v(r, \alpha) = (1 - r^2 \sin^2 \alpha)^{\frac{1}{2}}, \tag{5}$$

as seen from (3) and the conformity of the mapping $w = \log z$ (cf. Fig. 1 and Fig. 2 of Bohr and Jessen [4]). Hence, the expression on the right of (4) is

$$f_\sigma(\alpha) = 2^{-\sigma}(1 - 2^{-2\sigma} \sin^2 \alpha)^{-\frac{1}{2}} - \sum_{n=2}^{\infty} p_n^{-\sigma}(1 - p_n^{-2\sigma} \sin^2 \alpha)^{-\frac{1}{2}}, \tag{6}$$

so that, on differentiating with respect to α,

$$f'_\sigma(\alpha) = \sin\alpha\cos\alpha[2^{-3\sigma}(1-2^{-2\sigma}\sin^2\alpha)^{-(3/2)} - \sum_{n=2}^{\infty} p_n^{-3\sigma}(1-2^{-2\sigma}\sin^2\alpha)^{-(3/2)}].$$

Since if $a_n > 0$, then $a_1 \geqq \sum_{n=2}^{\infty} a_n$ implies $a_1^3 \geqq (\sum_{n=2}^{\infty} a_n)^3 > \sum_{n=2}^{\infty} a_n^3$, it follows that $f_\sigma(\alpha) \geqq 0$ implies $f'_\sigma(\alpha) > 0$. Hence, the expression on the right of (4) is a monotone increasing function of α for all those σ for which $f_\sigma(+0) \geqq 0$. Consequently, the radius of curvature $\rho_\sigma(\alpha)$ exists and is represented by (4) for every positive $\alpha \leqq \frac{1}{2}\pi$, if and only if σ is such that the expression on the right of (4) tends to a non-negative limit as $\alpha \to +0$. Now it will be shown that there exists a $\sigma^0 > 1$ such that $f_\sigma(+0)$ is negative, zero or positive according as $1 < \sigma < \sigma^0$, $\sigma = \sigma^0$ or $\sigma^0 < \sigma < +\infty$. First, it is seen from (6) that

$$f_\sigma(+0) = 2^{-\sigma} - \sum_{n=2}^{\infty} p_n^{-\sigma},$$

so that $f_\sigma(+0)$ is positive for large σ, negative for σ near enough to 1 and zero if and only if σ is a root of the equation (2b). Thus, it is sufficient to show that the equation

$$g(\sigma) = 1, \quad \text{where} \quad g(\sigma) = \sum_{n=2}^{\infty} (2/p_n)^\sigma$$

has exactly one root $\sigma = \sigma^0 > 1$. But this is obvious, since $g(\sigma)$ is a continuous, monotone decreasing function for which $g(1+0) = +\infty$, $g(+\infty) = 0$.

In order to complete the proof of (II), it remains to be shown that if $B(\sigma)$ has corners, they lie on the real axis. Suppose, if possible, that $B(\sigma)$ has, for some fixed $\sigma > 1$, a corner not on the real axis. Then it is seen from the double symmetry of $B(\sigma)$ that the α-interval $0 < \alpha \leqq \frac{1}{2}\pi$ contains a subinterval which corresponds to a corner. Let $\alpha_1 \leqq \alpha \leqq \alpha_2$ be this subinterval, so that $0 < \alpha_1 < \alpha_2 \leqq \frac{1}{2}\pi$. It is clear from the considerations of the paper [8], especially from the geometrical construction of $B(\sigma)$ given there, that one can choose this subinterval in such a way that $f_\sigma(\alpha)$ is negative for $\alpha_1 < \alpha < \alpha_2$ and zero for $\alpha = \alpha_1$. Hence $f_\sigma(\alpha)$ is certainly not increasing in the interval $\alpha_1 < \alpha \leqq \frac{1}{2}\pi$. On the other hand, it has been shown above that if $f_\sigma(\alpha_1) = 0$, then $f_\sigma(\alpha)$ is increasing for $\alpha_1 < \alpha \leqq \frac{1}{2}\pi$. This contradiction completes the proof of (II).

Incidentally, the above considerations clearly imply the following result:

(III) *The radius of curvature $\rho_\sigma(\alpha)$ of $B(\sigma)$ is, for every fixed $\sigma > \sigma_0$, defined, continuous and increasing in the α-interval $\alpha_\sigma(+0) < \alpha \leqq \frac{1}{2}\pi$. If $\sigma = \sigma^0$, then $\alpha_\sigma(+0) = 0$ and $\rho_\sigma(+0) = 0$ so that, while the convex curve*

$B(\sigma^0)$ has no corners, its radius of curvature vanishes at the points on the real axis.

If $\sigma \geqq \sigma^0$, then $B(\sigma)$ has no corners by (II), while $B(\sigma)$ does not exist for $\sigma \leqq \sigma_0$ by (I). The following theorem describes the situation in the remaining σ-interval.

(IV) *If $\sigma_0 < \sigma < \sigma^0$ and $\theta = \theta(\sigma)$ denotes the positive angle $(<\pi)$ formed by the two branches of $B(\sigma)$ at a corner of $B(\sigma)$, then $\theta(\sigma)$ is a monotone function of σ and $\theta(\sigma_0 + 0) = 0$, $\theta(\sigma^0 - 0) = \pi$.*

First, it is clear from the definition of $\alpha_\sigma(v)$, as given in the proof of (II), that

$$\theta(\sigma) = \pi - 2\alpha_\sigma(+0). \tag{7}$$

Furthermore, it is clear from the proof of Theorem II_0 in the paper [8] that if Q_a and $P_a{}^n$ denote the points of $B(\sigma)$ and $S(p_n{}^{-\sigma})$ which have the normal inclination α, then Q_a is the vectorial sum of the points $P^1{}_a, P^2{}_{a+\pi}, P^3{}_{a+\pi}, \cdots$. This implies, in view of the definitions of $v_\sigma(\alpha)$ and $v(r, \alpha)$ as given in the proof of (II), that

$$v_\sigma(\alpha) = v(2^{-\sigma}, \alpha) + \sum_{n=2}^{\infty} v(p_n{}^{-\sigma}, \alpha + \pi).$$

Now $v(r, \alpha) = -v(r, \alpha + \pi)$ in view of the summetry of $S(r)$. Consequently, since $\sin v(r, \alpha) = r \sin \alpha$ by (5),

$$v_\sigma(\alpha) = \arcsin\,(2^{-\sigma} \sin \alpha) - \sum_{n=2}^{\infty} \arcsin\,(p_n{}^{-\sigma} \sin \alpha).$$

On the other hand it is clear from the definition of α and from the convexity of $B(\sigma)$ that $\alpha_\sigma(v)$ is increasing in the interval $0 < v \leqq \sum_{n=1}^{\infty} \arcsin p_n{}^{-\sigma}$ so that $\alpha_\sigma(+0)$ is the largest root $(< \frac{1}{2}\pi)$ of the equation $v_\sigma(\alpha) = 0$, i. e., of the equation

$$F(\sigma, \alpha) = 0, \tag{8}$$

where

$$F(\sigma, \alpha) = \sum_{n=2}^{\infty} \arcsin\,(p_n{}^{-\sigma} \sin \alpha)/\arcsin\,(2^{-\sigma} \sin \alpha) - 1. \tag{9}$$

First, it will be shown by an argument very similar to that of Bohr and Jessen [4], pp. 42-43, that

$$\partial F(\sigma, \alpha)/\partial \alpha < 0, \qquad 0 < \alpha \leqq \tfrac{1}{2}\pi. \tag{10}$$

In fact, it will be shown that each term

$$\phi_n(\sigma, \alpha) = \arcsin\,(p_n^{-\sigma} \sin \alpha)/\arcsin\,(2^{-\sigma} \sin \alpha)$$

of (9) has a negative derivative, $\partial\phi_n/\partial\alpha$. Since the logarithmic derivative $\phi_n^{-1}\partial\phi_n/\partial\alpha$ of ϕ_n is the difference between the logarithmic derivatives of the numerator and the denominator, and since $\phi_n > 0$, in order to show that $\partial\phi_n/\partial\alpha < 0$ it is sufficient to prove that the logarithmic derivative $\psi^{-1}\partial\psi/\partial\alpha$ of the function $\psi(\alpha, \lambda) = \arcsin\,(\lambda \sin \alpha)$ is, for each fixed α in the interval $0 < \alpha \leqq \frac{1}{2}\pi$, an increasing function of λ in $0 < \lambda < 1$. Now the latter logarithmic derivative is

$$(11) \qquad \frac{\lambda \cos \alpha}{\sqrt{1 - \lambda^2 \sin^2 \alpha}\,\arcsin\,(\lambda \sin \alpha)}, \quad \text{i. e.,} \quad \cot \alpha \frac{\tan y}{y},$$

where $y = \arcsin\,(\lambda \sin \alpha)$, so that y is an increasing function of λ. Since $(\tan y)/y$ is an increasing function of y in $0 < y \leqq \frac{1}{2}\pi$ it follows that the logarithmic derivative (11) is an increasing function of λ. This completes the proof of (10).

Now it has been pointed out above that $\alpha = \alpha_\sigma(+0)$ is a root of the equation (8). On the other hand, $\alpha = \alpha_\sigma(+0)$ lies in the α-range admitted in (10), since $\alpha_\sigma(+0) = 0$ would imply that $B(\sigma)$ has no corners, whereas $\sigma_0 < \sigma < \sigma^0$ by assumption. On substituting $\alpha = \alpha_\sigma(+0)$ into (8), it follows, from (10) and from the obvious existence of continuous partial derivatives of the function (9), that the function $\alpha_\sigma(+0)$ of σ possesses the continuous derivative

$$(12) \quad d\alpha_\sigma(+0)/d\sigma = -\partial F(\sigma, \alpha)/\partial\sigma : \partial F(\sigma, \alpha)/\partial\alpha \quad (\alpha = \alpha_\sigma(+0)).$$

Now it is easily seen from the considerations of Bohr and Jessen ([4], pp. 42-43) that

$$(13) \qquad \partial F(\sigma, \alpha)/\partial\sigma < 0 \qquad (0 < \alpha \leqq \tfrac{1}{2}\pi,\ \sigma_0 < \sigma < \sigma^0).$$

On comparing (12) with (10) and (13) it follows that $d\alpha_\sigma(+0)/d\sigma < 0$, so that $\theta(\sigma)$ is, in view of (7) an increasing function of σ in the interval $\sigma_0 < \sigma < \sigma^0$. In order to complete the proof of (IV) it is sufficient to notice that $\theta(\sigma_0 + 0) = 0$ is obvious, if one compares the equation (2a) which defines σ_0 with the relations (8), (9), while $\theta(\sigma^0 - 0) = \pi$ similarly follows from the equation (2b) which defines σ^0, if one lets $\alpha \to 0$ in (9).

With regard to the shape of the region $M(\sigma)$ for large values of σ one may say the following:

(V) *The region $M(\sigma)$ is, for large values of σ, asymptotically a circle of radius $2^{-\sigma}$.*

In fact, let $h(r, \alpha)$ be the supporting function of $S(r)$ so that $h(r, \alpha) = u \cos \alpha + v \sin \alpha$, where (u, v) is the point of $S(r)$ for which the normal inclination has the value α. Then on introducing for u and v their expressions in terms of r and α and expanding in powers of r, it may easily be shown that

$$h(r, \alpha) \sim r \text{ as } r \to 0,$$

uniformly in α. Now if $g(\sigma, \alpha)$ is the supporting function of $A(\sigma)$, then, according to the rule of Haviland ([5], p. 333), $g(\sigma, \alpha) = \sum_{n=1}^{\infty} h(p_n^{-\sigma}, \alpha)$, so that

$$g(\sigma, \alpha) \sim 2^{-\sigma} + \sum_{n=2}^{\infty} p_n^{-\sigma} \sim 2^{-\sigma}; \qquad \sigma \to +\infty; \tag{14}$$

uniformly for all α.

Similarly, if $k(\sigma, \alpha)$ is the supporting function of $B(\sigma)$ and if $\sigma > \sigma^0$, then, by [8], p. 743, $k(\sigma, \alpha) = h(2^{-\sigma}, \alpha) - h(p_n^{-\sigma}, \alpha + \pi)$, so that

$$k(\sigma, \alpha) \sim 2^{-\sigma} - \sum_{n=2}^{\infty} p_n^{-\sigma} \sim 2^{-\sigma}, \qquad \sigma \to +\infty; \tag{15}$$

uniformly for all α. Since (14) and (15) hold uniformly for all α, the result (V) follows.

In conclusion, the following obvious remark concerning the shape of $B(\sigma)$, as $\sigma \to \sigma_0$, may be made.

(VI) *The ratio $R(\sigma)$ of the diameter of $B(\sigma)$ along the u-axis to the diameter parallel to the v-axis tends to infinity as $\sigma \to \sigma_0$.*

In fact, a trivial application of the mean value theorem gives $R(\sigma) > \tan \alpha_\sigma(+0)$, while $\tan \alpha_\sigma(+0) \to +\infty$ as $\sigma \to \sigma_0$, by (IV) and (7).

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] H. Bohr, "Sur la fonction $\zeta(s)$ dans le demi-plan $\sigma > 1$," *Comptes Rendus*, vol. 154 (1912), pp. 1078-1081.

[2] H. Bohr, "Om Addition af uendelig mange konvekse Kurver," *Danske Videnskabernes Selskab, Forhandlinger*, 1913, pp. 325-366.

[3] H. Bohr and B. Jessen, "Om Sandsynlighedsfordelinger ved Addition af konvekse Kurver," *Danske Videnskabernes Selskab*, Skrifter, (8), vol. 12, no. 3 (1929).

[4] H. Bohr and B. Jessen, "On the distribution of the values of the Riemann zeta function," *American Journal of Mathematics*, vol. 58 (1936), pp. 35-45.

[5] E. K. Haviland, "On the addition of convex curves in Bohr's theory of Dirichlet series," *American Journal of Mathematics*, vol. 55 (1933), pp. 332-334.

[6] B. Jessen and A. Wintner, "Distribution functions and the Riemann zeta function," *Transactions of the American Mathematical Society*, vol. 38 (1935), pp. 48-88.

[7] R. Kershner and A. Wintner, "On the boundary of the range of values of $\zeta(s)$," *American Journal of Mathematics*, vol. 58 (1936), pp. 421-426.

[8] R. Kershner, "On the addition of convex curves," *American Journal of Mathematics*, vol. 58 (1936), pp. 737-747. Formula (ii), p. 745, should read

$$\sqrt{A(C_1)} \leqq \sqrt{A(C_1)} - \sum_{j=2}^{n} \sqrt{A(C_j)}.$$

CONVOLUTIONS OF DISTRIBUTIONS ON CONVEX CURVES AND THE RIEMANN ZETA FUNCTION.*

By E. R. VAN KAMPEN and AUREL WINTNER.

Introduction. It is known that the almost periodic function

$$x + iy = z = z(t) = \log \zeta(\sigma + it), \qquad -\infty < t < +\infty, \tag{1}$$

where $\sigma > 1$ is fixed, has an absolutely continuous asymptotic distribution function the density of which possesses continuous partial derivatives of arbitrarily high order.[1] The same holds for this density $\delta = \delta^\sigma = \delta^\sigma(x, y)$ in the case $\sigma = 1$, while if $\frac{1}{2} < \sigma < 1$, then $\delta^\sigma(x, y)$ is a transcendental entire function of two variables.[2] These results are proved by estimating the order of magnitude of the Fourier transform in the infinity.[3] It is not known whether or not $\delta^\sigma(x, y)$ is, in the case $\sigma = 1$, regular analytic in the whole real (x, y)-plane, or at least in some portions of it. In what follows, it will be assumed that $\sigma > 1$. In this case the spectrum of the asymptotic distribution function of (1) is a bounded set, so that $\delta^\sigma(x, y)$, being identically zero without the spectrum, vanishes with all its derivatives on the boundary of the spectrum, and so $\delta^\sigma(x, y)$ cannot be regular analytic there. It may be mentioned that the boundary of the spectrum consists of regular analytic curves.[4] The question of subregions of analyticity within the spectrum cannot be discussed by the method of Fourier transforms mentioned above. In fact, this method yields the regular analyticity of $\delta^\sigma(x, y)$ either for every (x, y) or for no (x, y).

In what follows, there will be delimited within the spectrum subregions of regular analyticity for the density $\delta^\sigma = \delta^\sigma(x, y)$. It turns out that there exists a sequence $\alpha_1, \alpha_2, \cdots$ of numbers $\alpha_k > 1$ such that $\delta^\sigma(x, y)$ is for every $\sigma > \alpha_k$ regular analytic in a certain number, say N_k, of mutually disjoint ring-shaped subregions of the spectrum, and that $N_k \to +\infty$ as $k \to +\infty$. Since the N_k rings form a disconnected set, it cannot be stated that the corresponding N_k regular functions of two variables x, y are analytic continuations of each

* Received November 16, 1936.

[1] Wintner [16], pp. 328-329; Jessen and Wintner [10], Theorems 19 and 30; Haviland and Wintner [9].

[2] Jessen and Wintner [10], Theorems 19 and 30.

[3] Wintner [17]; Jessen and Wintner [10], p. 54.

[4] Kershner and Wintner [15].

other. There are some indications to the effect that the Jordan curves which form the boundary of the ring-shaped regions to be determined consist of singular points of $\delta^\sigma(x, y)$; but we cannot prove this. It is not stated that the N_k rings to be determined contain all subregions of analyticity within the spectrum.

The results in the Appendix concern the existence and a discussion of singularities within the spectrum of the asymptotic distribution functions of real almost periodic functions with linearly independent frequencies.

The proof to be given for the existence of rings of analyticity of $\delta^\sigma(x, y)$ for $\sigma > \alpha_k$ depends largely on the geometry of convolutions of distributions on convex curves, so that the essential part of the paper consists of necessary refinements of certain elementary geometrical results contained in the literature.[5] In fact, the results to be obtained on the geometry of the convolutions in question will make it possible to establish the existence of rings of analyticity by means of an argument previously applied [6] in the case of almost periodic functions with linearly independent frequencies, i. e., in the case of infinite convolutions of circular equidistributions. It may be mentioned that for the convergence of the infinite convolutions in question there exists, in case of a bounded spectrum, an entirely geometrical criterion, involving not the distributions along the partial convex curves but merely the latter curves themselves (Section 4).

On applying a trivial exponential mapping, the results obtained for the asymptotic distribution function of (1) yield corresponding results for the asymptotic distribution function of

$$x + iy = z = z(t) = \zeta(\sigma + it), \qquad -\infty < t < +\infty.$$

The method used in the case of $\log \zeta(s)$ will be applied to $\zeta'(s)/\zeta(s)$ also. While in the latter case the geometry of the problem [7] is as simple as in the case of circular equidistributions, the problem of the densities is more involved.

1. *The convolution of two distributions on convex curves.* Let Θ_1, Θ_2 be two oriented circles and let θ_j denote the angular coördinate on Θ_j. If A_1, A_2 are subsets of Θ_1, Θ_2 respectively, let $A_1 \times A_2$ denote the product space of A_1 and A_2, so that, in particular, $\Theta_1 \times \Theta_2$ denotes a torus on which a point is characterized by a pair (θ_1, θ_2) of angular coördinates. If A, A_j, where $j = 1, 2$, are Borel sets on $\Theta_1 \times \Theta_2$, Θ_j respectively, put

[5] Bohr [2], [3]; Bohr and Jessen [5]; Haviland [8]; Bohr and Jessen [6]; Kershner [11], [12].

[6] Kershner and Wintner [14].

[7] Bohr [1], Burrau [7].

$$(2) \qquad \mu(A) = \frac{\text{meas } A}{\text{meas } \Theta_1 \times \Theta_2}, \qquad \lambda(A_j) = \frac{\text{meas } A_j}{\text{meas } \Theta_j}, \qquad (j = 1, 2),$$

where the measures refer to those on $\Theta_1 \times \Theta_2$, Θ_j respectively, when $\Theta_1 \times \Theta_2$, Θ_j are thought of as represented by means of the parameters (θ_1, θ_2), θ_j. Thus $\mu(A)$, $\lambda(A_j)$ are non-negative absolutely additive set-functions which are defined for all Borel sets A, A_j on $\Theta_1 \times \Theta_2$, Θ_j respectively and have the total variation 1. The distribution functions $\mu(A)$, $\lambda(A_j)$ represent equidistributions with constant densities $(\frac{1}{2\pi})^2$, $\frac{1}{2\pi}$ on $\Theta_1 \times \Theta_2$, Θ_j respectively. It is clear that if A_1, A_2 are Borel sets on Θ_1, Θ_2 respectively, then the product space $A_1 \times A_2$ is a Borel set on the torus $\Theta_1 \times \Theta_2$ and

$$\mu(A_1 \times A_2) = \lambda(A_1)\lambda(A_2).$$

Let a function

$$(3) \qquad x + iy = z = z_j(\theta_j) \equiv \xi_j(\theta_j) + i\eta_j(\theta_j)$$

defined on Θ_j, where $j = 1, 2$, be an admissible parameter representation of a convex curve S_j in the z-plane, the adjective "admissible" being meant in the usual sense. In other words, on denoting by T_j the transformation (3) of the circle Θ_j into the convex curve S_j, it is assumed that T_j is a topological transformation and that the function (3) of θ_j has a continuous derivative which is nowhere zero:

$$(4) \qquad |z'_j(\theta_j)|^2 = \xi'_j(\theta_j)^2 + \eta'_j(\theta_j)^2 \neq 0.$$

It may be assumed that the orientation of Θ_j, when transplanted by means of T_j onto S_j, determines on S_j that orientation which is positive in the z-plane, so that the positively oriented normal at a point of S_j is the exterior normal at that point.

In addition to the assumption that (3) is an admissible parameter representation of the convex curve S_j, let it be supposed that there exists for every $\omega = \arg z$ exactly one point θ_j on Θ_j, say

$$(5) \qquad \theta_j = \vartheta_j(\omega), \qquad (j = 1, 2),$$

such that the positive normal at the point $z = z_j(\theta_j)$ of S_j has the normal inclination ω. In other words, it will be supposed that the convex curve S_j does not contain rectilinear segments. It is clear that (5) is an orientation-preserving topological transformation of the unit circle $z = e^{i\omega}$ of the z-plane into the circle Θ_j.

For a given Borel set E in the z-plane, let $\phi_j(E)$ denote $\lambda(T_j^{-1}(ES_j))$,

i. e., the value of the set-function $\lambda(\mathfrak{A}_j)$ in (2) for the Borel set $\mathfrak{A}_j = T_j^{-1}(ES_j)$ into which the inverse of the transformation (3) transforms the common part of the Borel set E and of the convex curve S_j. Thus $\phi_j(E)$ is a continuous, but not absolutely continuous, distribution function which has S_j as spectrum and is defined for all Borel sets E in the z-plane. The convolution [8] of the distribution functions ϕ_1, ϕ_2 will be denoted by ψ_2:

$$\psi_2(E) = \psi_2 = \phi_1 * \phi_2. \tag{6}$$

For a given pair of functions (3), let T denote the transformation

$$x + iy = z(\theta_1, \theta_2) \equiv \sum_{j=1}^{2} z_j(\theta_j) \equiv \sum_{j=1}^{2} \xi_j(\theta_j) + i \sum_{j=1}^{2} \eta_j(\theta_j) \tag{7}$$

of the torus $\Theta_1 \times \Theta_2$ into a subset of the z-plane. This subset, $T(\Theta_1 \times \Theta_2)$, of the z-plane is what is called the vectorial sum of the convex curves S_1, S_2 and will be denoted, in contradistinction to the logical sum $S_1 + S_2$, by $S_1(+)S_2$. The Jacobian of the transformation T is, according to (7),

$$J = J(\theta_1, \theta_2) = \frac{\partial(x, y)}{\partial(\theta_1, \theta_2)} = \begin{vmatrix} x'_1(\theta_1) & x'_2(\theta_2) \\ y'_1(\theta_1) & y'_2(\theta_2) \end{vmatrix}. \tag{8}$$

Now [9] the convolution (6) is in the same way the transplantation of the equidistribution $\mu(A)$ on $\Theta_1 \times \Theta_2$ onto the z-plane under the mapping T as the distribution function $\phi_j = \phi_j(E)$ is the transplantation of the equidistribution $\lambda(\mathfrak{A}_j)$ on Θ_j onto the z-plane under the mapping T_j, where $j = 1, 2$. More explicitly, $\psi_2(E)$ is, for every Borel set E in the z-plane, equal to $\mu(T^{-1}(E))$, where the Borel set $\mathfrak{A} = T^{-1}(E)$ is meant to be the inverse image of E under the transformation T, i. e., the set of those points (θ_1, θ_2) of $\Theta_1 \times \Theta_2$ which are transformed by the continuous, but not topological, transformation (7) into points z of E. In particular, $S_1(+)\, S_2$ is the spectrum of the distribution function (6). It is also seen that if

$$\bar{\Lambda} = \Theta_1 \times \Theta_2 - \sum_{k=1}^{m} \Lambda_k,$$

where $\Lambda_1, \cdots, \Lambda_m$ are mutually disjoint Borel sets on $\Theta_1 \times \Theta_2$, then

$$\psi_2(E) = \psi_2(E; \bar{\Lambda}) + \sum_{k=1}^{m} \psi_2(E; \Lambda_k), \tag{9a}$$

where

$$\psi_2(E; \Lambda) = \mu(\mathfrak{A}\Lambda), \qquad \mathfrak{A} = T^{-1}(E), \tag{9b}$$

[8] As to terminology and notations, cf. Jessen and Wintner [10], Section 2.

[9] Jessen and Wintner [10], p. 84.

so that $\psi_2(E;\Lambda)$ denotes the contribution of a portion Λ of $\Theta_1 \times \Theta_2$ to $\psi_2(E) = \psi_2(E;\Theta_1 \times \Theta_2)$.

2. *The set Ω and its complement on the torus.* Let Ω denote the set of those points (θ_1, θ_2) of the torus $\Theta_1 \times \Theta_2$ at which the Jacobian (8) of the transformation T vanishes. Thus the continuous transformation T defined by (7) is locally topological on the set

$$\bar{\Omega} = \Theta_1 \times \Theta_2 - \Omega,$$

which, Ω being obviously closed, is an open subset of $\Theta_1 \times \Theta_2$.

In order to describe Ω in terms of the two periodic functions (5), let there be defined on the torus $\Theta_1 \times \Theta_2$ two disjoint rectifiable Jordan curves Ω^+, Ω^- by means of the parameter representations

$$\Omega^+ : \theta_1 = \vartheta_1(\omega), \quad \theta_2 = \vartheta_2(\omega); \tag{10a}$$

$$\Omega^- : \theta_1 = \vartheta_1(\omega), \quad \theta_2 = \vartheta_2(\omega) + \pi. \tag{10b}$$

Then Ω is the logical sum

$$\Omega = \Omega^+ + \Omega^-. \tag{11}$$

In fact, it is clear from (4) and (8) that $J(\theta_1, \theta_2)$ vanishes at the point (θ_1, θ_2) of $\Theta_1 \times \Theta_2$ if and only if the oriented normal to S_1 at the T_1-image of the point θ_1 of Θ_1 is either parallel or anti-parallel to the oriented normal to S_2 at the T_2-image of the point θ_2 of Θ_2. This means in view of the definition of the functions (5) that $J(\theta_1, \theta_2) = 0$, i. e. $(\theta_1, \theta_2) \subset \Omega$, holds if and only if (θ_1, θ_2) is a point either of the curve (10a) or of the curve (10b). That the components Ω^+, Ω^- of the set (11) are disjoint rectifiable Jordan curves, is seen from (10a), (10b) and from the fact that each of the functions (5) defines, by Section 1, a topological transformation of a circle into another circle.

Let R denote the portion of $S_1(+)\, S_2$ not contained in the T-image of (11):

$$R = [S_1\ (+)\ S_2] - T(\Omega). \tag{12}$$

It is easy to see that R is an open set. In fact, $\bar{\Omega}$ is an open set on the torus $\Theta_1 \times \Theta_2$ and T is locally topological on $\bar{\Omega}$, so that $T(\bar{\Omega})$ is an open set in the z-plane. On the other hand, Ω is a closed subset of the compact set $\Theta_1 \times \Theta_2$ on which T is continuous, so that $T(\Omega)$ is a closed set in the z-plane. Now

$$S_1\ (+)\ S_2 = T(\Theta_1 \times \Theta_2) = T(\Omega + \bar{\Omega}) = T(\Omega) + T(\bar{\Omega}),$$

so that R is, in view of (12), the set of those points z which are in $T(\bar{\Omega})$ but not in $T(\Omega)$. Since $T(\bar{\Omega})$ is open and $T(\Omega)$ is closed, it follows that R is open.

Next, if $z = z_0$ is any fixed point of R and $\epsilon = \epsilon(z_0) > 0$ is so small that the circle

$$U : |z - z_0| \leqq \epsilon \tag{13a}$$

is contained in R, then there exists on the torus $\Theta_1 \times \Theta_2$ a finite number, say $m = m(U)$, of mutually disjoint closed sets

$$\Lambda_1, \cdots, \Lambda_k, \cdots, \Lambda_m \tag{13b}$$

such that every Λ_k is transformed by T into U in a topological way, while no point of

$$\bar{\Lambda} = \Theta_1 \times \Theta_2 - \sum_{k=1}^{m} \Lambda_k \tag{13c}$$

is transformed by T into a point of U. In fact, since T is locally topological on $\bar{\Omega}$ and R was seen to be a subset of $T(\bar{\Omega})$, it is clear from the monodromy theorem that it is sufficient to prove that the T^{-1}-image of any fixed point z_0 of R is a finite subset of $\bar{\Omega}$. Now suppose, if possible, that the T^{-1}-image of a z_0 is not a finite set and has, therefore, a cluster point ζ on the torus $\Theta_1 \times \Theta_2$. Since T is locally topological on $\bar{\Omega}$, the cluster point ζ cannot be a point of $\bar{\Omega}$, so that ζ must be a point of the complement Ω of $\bar{\Omega}$. This is, however, impossible, since the T^{-1}-image of the point z_0 is a closed set contained in $\bar{\Omega}$.

It has been shown by Bohr [10] that the vectorial sum $S_1 (+) S_2$ mentioned in Section 1, i. e. the T-image of the torus $\Theta_1 \times \Theta_2$, is either a closed bounded convex region or a closed bounded ring-shaped region bordered by two convex curves which have no point in common. In the first case let C_2 denote the convex curve which forms the boundary of $S_1 (+) S_2$; in the second case let C_2 and D_2 denote the convex curves which form the outer and inner boundary of $S_1 (+) S_2$ respectively. Now the convex curve C_2 (in both cases) and the convex curve D_2 (in the second case) are related to the curves Ω^+ and Ω^- on the torus $\Theta_1 \times \Theta_2$ as follows:

$$T(\Omega^+) = C_2, \qquad T^{-1}(C_2) = \Omega^+,$$

and, if D_2 exists,

$$T(\Omega^-) \supset D_2, \qquad T^{-1}(D_2) \subset \Omega^-,$$

where

$$T(\Omega^-) = D_2, \qquad T^{-1}(D_2) = \Omega^-$$

[10] Bohr [2]; cf. also the presentation in Jessen and Wintner [10], Section 9.

holds if and only if D_2 is free of corners. These facts may be obtained by comparing the above considerations with those of Haviland [11] and Kershner [12] with regard to C_2 and D_2 respectively. Hence if D_2 exists and is free of corners, then the open set R defined by (12) and (11) is identical with the open ring-shaped region bordered by the boundary curves C_2, D_2 of the vectorial sum $S_1 (+) S_2$.

3. *The density of the convolution* (6). In this Section 3 there will be obtained for the distribution function ψ_2 defined by (6) an explicit representation in terms of the Jacobian (8). This representation will not only imply that ψ_2 is, in contrast with ϕ_1 and ϕ_2, an absolutely continuous distribution function but will also show that the density $\delta_2 = \delta_2(x, y)$ of ψ_2 is always continuous on the open set R, finally that $\delta_2(x, y)$ is regular analytic on R in case the functions (3) are regular analytic on Θ_1, Θ_2. Incidentally, $\delta_2(x, y)$ will be seen to be non-bounded on any open set containing a boundary point of R. A proof of the absolute continuity of ψ_2 has been indicated by Bohr.[13] The sharper results to be obtained, which are indispensable for the final purpose of the present paper, will be proved by considerations suggested by Bohr's remarks but will be more elementary in nature, since no use need be made of Lebesgue integrals.

For any fixed point z_0 of the open set R, choose an $\epsilon > 0$ so small that the circle U defined by (13a) is contained in R. Then the T^{-1}-image of U consists, by Section 2, of the $m = m(U)$ closed sets (13b) each of which is transformed by T into U in a topological way. Let

$$\theta_1 = \alpha_k(x, y), \qquad \theta_2 = \beta_k(x, y), \qquad (k = 1, 2, \cdots, m), \tag{14}$$

denote the topological transformation T^{-1} of U into Λ_k. Since every Λ_k is a closed subset of $\bar{\Omega}$, it is clear from the definition of Ω given at the beginning of Section 2 that the reciprocal value of the Jacobian (8) is bounded on each of the m sets Λ_k. Finally, as pointed out in Section 1, the convolution (6) may be obtained by transplanting the distribution function $\mu(A)$ from the torus $\Theta_1 \times \Theta_2$ onto the z-plane by means of the transformation T; and $\mu(A)$ is the equidistribution on $\Theta_1 \times \Theta_2$, having the constant density $(\frac{1}{2\pi})^2$.

On comparing these facts with the definition (9b) of the symbol $\psi_2(E; \Lambda)$, it is clear from (14), (8) and from the transformation formula of Riemann double integrals that, for any rectangle E contained in the circle (13a),

[11] Haviland [8].
[12] Kershner [11].
[13] Bohr [3].

$$\psi_2(E;\Lambda_k) = \iint_E |\, 4\pi^2 J(\alpha_k(x,y),\beta_k(x,y))\,|^{-1}\,dx\,dy, \tag{15a}$$

where $k = 1, \cdots, m$. Similarly, since the T^{-1}-image of U consists of the closed sets (13a) and has, therefore, neither a point nor a cluster point in the set (13c), the application of the definition (9b) to the set $\Lambda = \bar{\Lambda}$ clearly gives

$$\psi_2(E;\bar{\Lambda}) = 0 \tag{15b}$$

for any rectangle E contained in the circle (13a). On substituting (15a), (15b) into (9a), it is seen that

$$\psi_2(E) = \iint_E \sum_{k=1}^{m} |\, 4\pi^2 J(\alpha_k(x,y),\beta_k(x,y))\,|^{-1}\,dx\,dy, \tag{16}$$

if E is a rectangle contained in the circle (13a). Since z_0 in (13a) is any point of R, it follows by writing z instead of z_0 that the set-function $\psi_2(E)$ is absolutely continuous on the subset R of $S_1\,(+)\,S_2$ and has at any point $z = x + iy$ of R the density

$$\delta_2(x,y) = \sum_{k=1}^{m} |\, 4\pi^2 J(\alpha_k(x,y),\beta_k(x,y))\,|^{-1}, \tag{17}$$

where the integer m or the functions (14) may vary if and only if one chooses the point $z = x + iy$ on distinct connected parts of the open set R. On the other hand, $\psi_2(E)$ is absolutely continuous on the complement of the spectrum $S_1\,(+)\,S_2$ with respect to the whole z-plane, the density $\delta_2(x,y)$ being identically zero on this complement. Hence it is clear from (12) that $\psi_2(E)$ is an absolutely continuous distribution function if and only if

$$\psi_2(T(\Omega)) = 0. \tag{18}$$

Now, while it is clear from (17) and from the definition of Ω that the density $\delta_2(x,y)$ is non-bounded in any vicinity of any point of the boundary $T(\Omega)$ of R, it is easy to see that (18) is true.

First, since $\psi_2(E)$ is the transplantation of the equidistribution $\mu(\mathbf{A})$ on $\Theta_1 \times \Theta_2$ onto the z-plane, (18) is clearly equivalent to

$$\mu(\mathbf{A}_0) = 0, \text{ where } \mathbf{A}_0 = T^{-1}(T(\Omega)). \tag{19}$$

Needless to say, one cannot write Ω instead of $T^{-1}(T(\Omega))$, since T^{-1} is multivalued on the T-image of Ω. Now Ω consists, according to (11), of two rectifiable Jordan curves on the torus $\Theta_1 \times \Theta_2$, so that, if $\epsilon > 0$ is given, one

can choose on $\Theta_1 \times \Theta_2$ an open set A_ϵ which contains Ω, consists of two ring-shaped regions bounded by rectifiable Jordan curves and is such that $\mu(A_\epsilon) < \epsilon$. Hence, in order to prove (19), it is sufficient to show that the common part $A_0\bar{A}_\epsilon$ of $A_0 = T^{-1}(T(\Omega))$ and of the complement $\bar{A}_\epsilon = \Theta_1 \times \Theta_2 - A_\epsilon$ of A_ϵ has a vanishing Jordan content on $\Theta_1 \times \Theta_2$. Since the reciprocal value of the Jacobian (8) is bounded on $\bar{A}_\epsilon$, it is thus sufficient to show that $T(\Omega)$ has a vanishing planar Jordan content. But this is obvious, since the Jacobian (8) is bounded on $\Theta_1 \times \Theta_2$ and Ω consists of two rectifiable Jordan curves on $\Theta_1 \times \Theta_2$.

The result thus proved may be formulated as follows: On defining in the z-plane a function $\delta_2(x, y) \geqq 0$ by placing it equal to the expression (17) or to 0 according as $z = x + iy$ is or is not in R, the convolution (6) may be represented for any Borel set E in the form

$$\psi_2(E) = \int_E \int \delta_2(x, y)\, dx\, dy. \tag{20}$$

Furthermore, if E is a closed rectangle, then (20) exists as a proper or as an improper Riemann integral according as E does or does not contain a boundary point of R. Incidentally, the above considerations imply that the closure of the set (12) is the spectrum $S_1\ (+)\ S_2$ of (6).

Remark. Suppose that the curves (3) are defined and satisfy the requirements of Section 1 not only for $j = 1, 2$ but for $j = 1, 2, \cdots, n$. Then if $\phi_1, \phi_2, \cdots, \phi_n$ denote corresponding distribution functions, the convolution $\phi_1 * \phi_2 * \cdots * \phi_n$ is absolutely continuous for every $n \geqq 2$ and has a continuous density for every $n \geqq 4$. In fact, the first statement is true for $n = 2$, and so it is true for every $n > 2$, since the convolution of an absolutely continuous and of an arbitrary distribution function is absolutely continuous.[14] For similar reasons, it is sufficient to show that the second statement, which has been proved by Bohr and Jessen [15] under somewhat more restrictive conditions and on using elementary geometrical considerations, is true in the lowest case, $n = 4$. Now if δ_2^{I} and δ_2^{II} denote the densities of $\phi_1 * \phi_2$ and $\phi_3 * \phi_4$ respectively, $\phi_1 * \phi_2 * \phi_3 * \phi_4$ is absolutely continuous with the density

$$D(x, y) = \int_{-\infty}^{+\infty} \int \delta_2^{\mathrm{I}}(x - \xi, y - \eta)\, \delta_2^{\mathrm{II}}(\xi, \eta)\, d\xi\, d\eta,$$

[14] Cf., e. g., Kershner and Wintner [13], p. 543, footnote.

[15] Bohr and Jessen [5].

so that, since $\delta_2^{\mathrm{I}}(\xi,\eta)$, $\delta_2^{\mathrm{II}}(\xi,\eta)$ are non-negative and have finite integrals $(=1)$ over the whole (ξ,η)-plane, $D(x,y)$ is everywhere continuous.

4. *Infinite convolutions.* Let $S_1, S_2, \cdots$ be an infinite sequence of convex curves S_m in the z-plane such that every S_m has a continuously turning tangent and does not contain rectilinear segments. Let

$$S_m : x+iy=z=z_m(\theta)\equiv\xi_m(\theta)+i\eta_m(\theta); \quad 0\leqq\theta<2\pi, \tag{21}$$

be an admisssible parameter representation of S_m in terms of an angular parameter θ and let $\phi_m=\phi_m(E)$ denote the corresponding distribution function defined for all Borel sets E of the z-plane (cf. Section 1). Thus S_m is the spectrum of ϕ_m, and so the spectrum of the convolution

$$\psi_n(E)=\psi_n=\psi_1*\cdots*\psi_n \tag{22}$$

is the vectorial sum

$$V_n=S_1\,(+)\cdots(+)\,S_n, \tag{23}$$

where $n=2,3,\cdots$.

Let d_m denote the diameter of S_m, the diameter of a set being meant in the usual sense. Suppose that

$$\sum_{m=1}^{\infty} d_m \text{ is convergent.}^{16} \tag{24}$$

Then the infinite convolution

$$\psi=\phi_1*\phi_2*\cdots \tag{25}$$

is convergent if and only if there exists on or within every S_m a point z_m such that

$$\sum_{m=1}^{\infty} z_m \tag{26}$$

is convergent, in which case (26) is uniformly convergent when every z_m varies on or within S_m. This clearly implies that if (24) is satisfied, then the infinite convolution (25) is absolutely convergent if and only if (26) is absolutely convergent for at least one choice of $z_1, z_2, \cdots$ on or within $S_1, S_2, \cdots$, in which case (26) is absolutely-uniformly convergent for all choices of $z_1, z_2, \cdots$ on or within $S_1, S_2, \cdots$. In particular, the infinite convolution (25) is absolutely convergent whenever (24) is satisfied and every S_m surrounds

[16] While the direction of a maximal cord of S_m may vary with m, an elementary consideration shows that (24) is satisfied if and only if $v_n<\text{const.}$, where v_n denotes the diameter of the set (23).

the origin of the z-plane. The interest of these criteria lies in the fact that no mention is made of the parameter representations (21) which determine the distribution functions ϕ_m occurring in (25), i. e., that if (24) is satisfied, then the convergence or absolute convergence of (25) depends only on the curves $S_1, S_2, \cdots$ themselves.

First, on placing

$$\gamma_m = \frac{1}{2\pi} \int_0^{2\pi} z_m(\theta)\, d\theta \tag{27}$$

and noting that (21) is a parameter representation of the convex curve S_m, it is clear that, d_m being the diameter of S_m, every point $z = z_m$ on or within S_m satisfies the inequality

$$| z_m - \gamma_m | \leqq d_m. \tag{28}$$

Since the point $z = z_m(\theta)$ is on S_m, it follows that d_m^2 is an upper bound for the value of

$$\mu_m = \frac{1}{2\pi} \int_0^{2\pi} | z_m(\theta) - \gamma_m |^2 \, d\theta. \tag{29}$$

Hence the assumption (24) implies that

$$\sum_{m=1}^{\infty} \mu_m \text{ is convergent.} \tag{30}$$

It is also seen from (28) and (24) that (26) is uniformly convergent for all choices of $z_1, z_2, \cdots$ on or within $S_1, S_2, \cdots$ whenever (26) is convergent for a single choice of $z_1, z_2, \cdots$ on or within $S_1, S_2, \cdots$; for instance, whenever

$$\sum_{m=1}^{\infty} \gamma_m \tag{31}$$

is a convergent series (in fact, $z = \gamma_m$ is, by (27), a point within S_m). Now if (31) is convergent, then it is clear from (21), (28) and (24) that all S_m are contained in a sufficiently large circular disk, and so (30) and the convergence of (31) imply [17] the convergence of (25). Conversely, if (25) is convergent, then [18] the convex curve S_m tends, as $m \to +\infty$, to the point $z = 0$, so that all S_m are contained in a sufficiently large circular disk, and

[17] Jessen and Wintner [10], Theorem 5.

[18] *Ibid.*, Theorem 1 ($p = 1$).

so the convergence of (25) implies [19] the convergence of (31). This clearly completes the proof of the facts stated above.

In the cases to be considered in the following Sections, the infinite convolution (25) is convergent and is known [20] to be absolutely continuous with a density $\delta(x, y)$ which has continuous partial derivatives of arbitrarily high order, while the finite convolution (22) is absolutely continuous with a density $\delta_n(x, y)$ which has [21] continuous partial derivatives of order k for every $n \geqq 2k + 5$. This implies [22] that

$$\delta(x, y) = \lim_{n=\infty} \delta_n(x, y) \tag{32}$$

for every (x, y) in the z-plane. Finally, all the densities $\delta_n(x, y)$ occurring in (32) may be obtained from the density $\delta_2(x, y)$ discussed in Section 3 by means of the recursion formula

$$\delta_{n+1}(x, y) = \frac{1}{2\pi} \int_0^{2\pi} \delta_n(x - \xi_{n+1}(\theta), y - \eta_{n+1}(\theta))\, d\theta. \tag{33}$$

This is seen from (21) and from the definition of the corresponding distribution function $\phi_m(E)$, since $\psi_{n+1} = \psi_n * \phi_{n+1}$ by (22).

The spectrum (23) of (22) is [23] either a closed bounded region bordered by a convex curve C_n or a closed bounded ring-shaped region bordered by two disjoint convex curves C_n, D_n, where D_n will denote the inner boundary of (23). Similarly,[24] if (25) is convergent, then its spectrum is the infinite vectorial sum

$$V = S_1 (+) S_2 (+) \cdots \tag{34}$$

which, when a bounded set, is either a closed convex region bordered by a convex curve C or a ring-shaped region bordered by two disjoint convex curves C, D, where D will denote the inner boundary of (34). The rôle of the assumption (24) made in the above geometrical criterion for the convergence of (25) is that if (25) is convergent, then its spectrum (34) is a bounded

---

[19] *Ibid.*, Theorem 5.

[20] Wintner [16], pp. 328-329; Jessen and Wintner [10], Sections 7 and 13 and p. 53; Haviland and Wintner [9].

[21] *Ibid.*

[22] Wintner [18].

[23] Bohr [2]; Jessen and Wintner [10], Theorem 17.

[24] Cf. Bohr [2]; Jessen and Wintner [10], Theorems 3 and 17.

set [25] if and only if the diameters d_m of the spectra S_m of the distribution functions ϕ_m satisfy the condition (24). This condition for the boundedness of (34) will be satisfied in the cases to be considered, so that the density (32) of (25) cannot be regular analytic for every (x, y); cf. the Introduction.

In order to obtain, in the case belonging to (1), subregions of the spectrum (34) of (25) such that the density (32) of (25) is a regular analytic function of the two real variables x, y within each of these subregions, the results of Section 3 will be combined in Section 7 with the method previously applied [26] to the case of circular equidistribution, i. e., to the case which belongs to almost periodic functions with linearly independent frequencies. Since the proof is, in the case which belongs to (1), quite involved, first the known case of circular equidistributions will be treated by means of the direct method instead of the reduction to the one-dimensional case previously applied.[26]

5. *Almost periodic functions with linearly independent frequencies.* The asymptotic distribution function of the almost periodic function

$$f(t) = \sum_{m=1}^{\infty} a_m \exp i\lambda_m(t + \alpha_m), \text{ where } a_m > 0, \tag{35}$$

is [27] in case of linearly independent frequencies λ_m the infinite convolution (25), where $\phi_m(E)$ is the distribution function belonging to (21), if

$$S_m : x + iy = z = \xi_m(\theta) + i\eta_m(\theta) \equiv a_m \exp i\theta. \tag{36}$$

Since (35) is almost periodic with linearly independent frequencies, the series $\sum_{m=1}^{\infty} a_m$, where $a_m > 0$, is convergent by a well-known theorem of Bohr. Put

$$r_h = \sum_{m=h+1}^{\infty} a_m, \qquad (h = 0, 1, 2, \cdots), \tag{37}$$

and let, without loss of generality,

$$a_m \geqq a_{m+1} > 0. \tag{38}$$

Then the spectrum (34) of (25) is seen to be the circle $|z| = r_0$ or the circular ring $2a_1 - r_0 \leqq |z| \leqq r_0$ according as $2a_1 - r_0 = 0$ or $2a_1 - r_0 > 0$. The recursion formula (33) becomes

---

[25] This is clear by combining the remark of footnote [16] with Theorem 3 of Jessen and Wintner [10]. Notice that Theorem 18 given there assumes that the convex curves S_n surround the point $z = 0$.

[26] Kershner and Wintner [14].

[27] Jessen and Wintner [10], Section 13.

$$(39) \qquad \delta_{n+1}(x,y) = \frac{1}{2\pi}\int_0^{2\pi} \delta_n(x - a_{n+1}\cos\theta, y - a_{n+1}\sin\theta)\,d\theta$$

in view of (36).

It is easily verified that the open set R defined by (12) is, in the case where S_1, S_2 are defined by (36), the punctured circle

$$(40a) \qquad 0 < |z| < a_1 + a_2, \qquad (z = x + iy),$$

or the circular ring

$$(40b) \qquad a_1 - a_2 < |z| < a_1 + a_2, \qquad (z = x + iy),$$

according as $a_1 = a_2$ or $a_1 > a_2$, the sets $T(\Omega^+)$, $T(\Omega^-)$ being the sets $|z| = a_1 + a_2$, $|z| = a_1 - a_2$ in both cases. The representation (17) of the density $\delta_2(x,y)$ of (6) within R becomes

$$(40) \qquad \delta_2(x,y) = \{\pi^2[(a_1+a_2)^2 - (x^2+y^2)][(x^2+y^2) - (a_1-a_2)^2]\}^{-1}.$$

Now suppose that

$$(41_2) \qquad a_2 > r_2.$$

For a given positive $\epsilon < a_2 - r_2$, choose a positive $\epsilon = \bar{\epsilon}(\epsilon)$ so small that on replacing (x,y) in (40) by $(x+iu, y+iv)$, where u, v are real, one obtains a function

$$(42_2) \qquad \delta_2(x+iu, y+iv)$$

which is regular analytic and bounded in the open set

$$(43_2) \qquad G_2 : \begin{cases} a_1 - a_2 + \epsilon < (x^2+y^2)^{\frac{1}{2}} < a_1 + a_2 - \epsilon, \\ 0 \leqq (u^2+v^2)^{\frac{1}{2}} < \bar{\epsilon} \end{cases}$$

of the two complex variables $x+iu$, $y+iv$. Since $\epsilon < a_2 - r_2 > 0$, it is clear from the definition (37) of r_h that

$$(43_n) \quad G_n : \begin{cases} a_1 - a_2 + \epsilon + \sum_{m=3}^{n} a_m < (x^2+y^2)^{\frac{1}{2}} < a_1 + a_2 - \epsilon - \sum_{m=3}^{\infty} a_m, \\ 0 \leqq (u^2+v^2)^{\frac{1}{2}} < \bar{\epsilon} \end{cases}$$

defines, for every $n \geqq 3$, an non-empty open set in the space of the two complex variables $x+iu$, $y+iv$. Let ϵ, $\bar{\epsilon}$ be fixed and let M denote a bound of the function (42_2) in the domain (43_2). Now it is clear from definitions (43_2), (43_n) that if G_{n+1}, where $n \geqq 2$, contains the point

$$(x+iu, y+iv), \text{ then } (x+iu - a_{n+1}\cos\theta, y+iv - a_{n+1}\sin\theta)$$

is a point of G_n for every θ. Hence, if there has been defined on G_n a regular analytic function

$$(42_n) \qquad \delta_n(x+iu, y+iv)$$

of $x+iy$, $u+iv$ such that $|\delta_n| \leqq M$ in G_n, then the formula

$$(44) \quad \delta_{n+1}(x+iu, y+iv) = \frac{1}{2\pi}\int_0^{2\pi} \delta_n(x+iu-a_{n+1}\cos\theta, y+iv-a_{n+1}\sin\theta)\,d\theta$$

defines on G_{n+1} a regular analytic function which has, in G_{n+1}, an absolute value

$$|\delta_{n+1}| \leqq \frac{1}{2\pi}\int_0^{2\pi} M\,d\theta = M.$$

Now if $u=0$, $v=0$, then (44) goes over into the representation (39) of the density $\delta_{n+1}(x, y)$ of ψ_{n+1}. Furthermore, the four-dimensional open set

$$(43_\infty) \quad G_\infty : \begin{cases} a_1 - a_2 + \epsilon + r_2 < (x^2+y^2)^{\frac{1}{2}} < a_1 + a_2 - \epsilon - r_2, \\ 0 \leqq (u^2+v^2)^{\frac{1}{2}} < \bar{\epsilon} \end{cases}$$

where $0 < \epsilon < a_2 - r_2$ by assumption, is, in view of (37), clearly contained in every G_n, where $n \geqq 2$. Finally, (32) holds for every real (x, y), and so, in particular, at every real point $(u=0, v=0)$ of the domain G_∞ defined by (43_∞). It follows, therefore, from the extension of Vitali's theorem to uniformly bounded sequences of regular functions of two variables that if $n \to +\infty$ and $(x+iu, y+iv)$ is in G_∞, then the functions (42_n) tend to a function $\delta(x+iu, y+iv)$ which is regular analytic in G_∞ and reduces, if $u=0$, $v=0$, to the density $\delta(x, y)$ of the infinite convolution (25). Since $\epsilon > 0$ in (43_∞) is arbitrarily small, it follows that the density (32) of (25) is a regular analytic function of the two real variables x, y in the ring

$$(45_2) \qquad a_1 - a_2 + r_2 < (x^2+y^2)^{\frac{1}{2}} < a_1 + a_2 - r_2,$$

whenever (41_2) is satisfied.

It is clear that if the assumption (41_2) is replaced by

$$(41_k) \qquad a_k > r_k,$$

where $k \geqq 2$ has a fixed value, an obvious modification of the above proof yields that the density (32) of (25) is a regular analytic function of the two real variables x, y in the ring

$$(45_k) \qquad \sum_{m=1}^{k-1} a_m - a_k + r_k < (x^2+y^2)^{\frac{1}{2}} < \sum_{m=1}^{k-1} a_m + a_k - r_k.$$

Furthermore, it is easily seen by the same argument that if (41_k) holds not only for a fixed $k \geqq 2$ but for $k = 1, \cdots, l$, where $l \geqq 2$ is fixed, then the density (32) of (25) is a regular analytic function of the two variables x, y not only within each of the $l-1$ rings defined by (45_k) for $k = 2, \cdots, l$, but also within each of the $2^{l-1}-1$ mutually disjoint rings

$$(46) \qquad \sum_{m=1}^{k-1} e_m a_m - a_k + \sum_{m=k+1}^{\infty} a_m < (x^2+y^2)^{\frac{1}{2}} < \sum_{m=1}^{k-1} e_m a_m + a_k - \sum_{m=k+1}^{\infty} a_m,$$

where $\sum_{m=k+1}^{\infty} a_m = r_k$ and

(47) $e_1 = 1$, while $e_m = \pm 1$ for $m = 2, 3, \cdots, k-1$; and $k = 2, 3, \cdots, l$,

it being understood that, since $e_1 = 1$, there are 2^{k-2} symbols $(e_1, \cdots, e_{k-1})$ for a fixed $k \geqq 2$.

Let it be mentioned for later reference that the a_m occurring in (46) are the supporting functions of the convex curves S_m, the latter being the circles $|z| = a_m$.

6. *The logarithmical derivative of the Riemann zeta function.* If $\sigma > 1$ is fixed, $s = \sigma + it$ and p_m is the m-th prime number, then the asymptotic distribution function of the almost periodic function

$$(48) \qquad f_\sigma(t) \equiv -\frac{\zeta'(s)}{\zeta(s)} + \text{const.} = \sum_{m=1}^{\infty} \frac{\log p_m}{p_m{}^s - 1} + \text{const.},$$

where const. $= \zeta'(2\sigma)/\zeta(2\sigma)$, may be represented [27] as the infinite convolution

$$(49) \qquad \psi^\sigma(E) = \psi^\sigma = \lim_{n=\infty} \psi_n{}^\sigma = \phi_1{}^\sigma * \phi_2{}^\sigma * \cdots,$$

where

$$(50) \qquad \psi_n{}^\sigma(E) = \psi_n{}^\sigma = \phi_1{}^\sigma * \phi_2{}^\sigma * \cdots * \phi_n{}^\sigma$$

and $\phi_m{}^\sigma = \phi_m{}^\sigma(E)$ denotes the distribution function which belongs to the curve (21) defined by

$$(51) \qquad S_m : x + iy = \xi_m(\theta) + i\eta_m(\theta) \equiv \frac{\log p_m}{p_m{}^\sigma e^{i\theta} - 1} - \frac{\log p_m}{p_m{}^{2\sigma} - 1}.$$

Thus S_m is the circle $|z| = a_m{}^\sigma$, where

$$(52) \qquad a_m{}^\sigma = \frac{p_m{}^\sigma \log p_m}{p_m{}^{2\sigma} - 1}, \qquad (a_m{}^\sigma > a_{m+1}{}^\sigma),$$

so that, since $\sigma > 1$, the series $\sum_{m=1}^{\infty} a_m{}^{\sigma}$ is convergent. Put

$$r_h{}^{\sigma} = \sum_{m=h+1}^{\infty} a_m{}^{\sigma}, \qquad (h = 0, 1, \cdots), \tag{53}$$

and let σ_k be the obviously unique number such that

$$a_k{}^{\sigma} = r_k{}^{\sigma} \text{ if } \sigma = \sigma_k > 1, \qquad (k = 1, 2, \cdots), \tag{54}$$

so that

$$a_k{}^{\sigma} > r_k{}^{\sigma} \text{ if and only if } \sigma > \sigma_k. \tag{55}$$

According to the rule (34), the spectrum of the infinite convolution (49) is the infinite vectorial sum

$$V^{\sigma} = S_1{}^{\sigma}(+)\, S_2{}^{\sigma}(+) \cdots. \tag{56}$$

It follows [28] that the spectrum of (49) is the circle $|z| \leqq r_0{}^{\sigma}$ or the circular ring $2a_1{}^{\sigma} - r_0{}^{\sigma} \leqq |z| \leqq r_0$ according as $\sigma \leqq \sigma_1$ or $\sigma > \sigma_1$, where $\sigma_1 = 2.576\cdots$ by (54).

Now let, for a fixed integer $l \geqq 2$,

$$\sigma > \text{Max}\,(\sigma_1, \cdots, \sigma_l). \tag{57}$$

Then [29] the density $\delta^{\sigma} = \delta^{\sigma}(x, y)$ of the distribution function (49) is a regular analytic function of the two real variables x, y within each of the $2^{l-1} - 1$ mutually disjoint rings which result from (46), (47) by writing $a_m{}^{\sigma}$ instead of a_m in (46).

In fact, the considerations of Section 5 may be repeated without any change if one proves that the density $\delta_2{}^{\sigma}(x, y)$ of $\phi_1{}^{\sigma} * \phi_2{}^{\sigma}$ is a regular analytic function of the two real variables x, y within the ring

$$a_1{}^{\sigma} - a_2{}^{\sigma} < (x^2 + y^2)^{\frac{1}{2}} < a_1{}^{\sigma} + a_2{}^{\sigma}.$$

Now this ring is, corresponding to (40b) and (55), identical with the open set R on which (17) is valid; and it is clear from (51) that (17) represents a regular analytic function on R. In fact, the functions (14) are regular analytic on R, since the Jacobian (8) vanishes only on Ω and the open set (12) does not contain points of the T-image of Ω.

---

[28] Bohr [1], Burrau [7].

[29] It will not be discusssed which of the l absolute constants $\sigma_1, \cdots, \sigma_l$ is Max $(\sigma_1, \cdots, \sigma_l)$ for a fixed l and what is the asymptotic behavior of σ_l and Max $(\sigma_1, \cdots, \sigma_l)$ for large l.

7. *The logarithm of the Riemann zeta function.* If $\sigma > 1$ is fixed, $s = \sigma + it$ and p_m is the m-th prime number, then the asymptotic distribution function of the almost periodic function

$$(58) \qquad f_\sigma(t) \equiv -\log \zeta(s) + \text{const.} = \sum_{m=1}^{\infty} \log(1 - p_m^{-s}) + \text{const.},$$

where const. $= \frac{1}{2} \log \zeta(2\sigma)$, may be represented [27] as the infinite convolution $\psi^\sigma(E)$ defined by (49) and (50), where $\phi_m{}^\sigma = \phi_m{}^\sigma(E)$ denotes the distribution function which belongs to the curve (21) defined by

$$(59) \quad S_m{}^\sigma : x + iy = \xi_m(\theta) + i\eta_m(\theta)$$
$$= \log(1 - p_m{}^{-\sigma} e^{i\theta}) - \tfrac{1}{2} \log(1 - p_m{}^{-2\sigma}) \equiv g(\theta; p_m{}^{-\sigma}) + ih(\theta; p_m{}^{-\sigma});$$
$$g(\theta;\rho) = \tfrac{1}{2} \log \frac{1 - 2\rho\cos\theta + \rho^2}{1 - \rho^2}, \qquad h(\theta;\rho) = \arctan \frac{\rho \sin\theta}{1 - \rho\cos\theta},$$
$$|h(\theta;\rho)| < \tfrac{1}{2}\pi.$$

Thus S_m is a regular analytic convex curve which has both axes $x = 0$, $y = 0$ as lines of symmetry.[30]

Let $a_m{}^\sigma(\omega)$ denote the supporting function (Stützfunktion) of $S_m{}^\sigma$, i. e., the distance of the origin $z = 0$ from that oriented normal of $S_m{}^\sigma$ which has the inclination ω. Then [30]

$$(60) \qquad \begin{aligned} a_m{}^\sigma(\tfrac{1}{2}\pi) &= \arcsin p_m{}^{-\sigma} (< \tfrac{1}{2}\pi), \\ a_m{}^\sigma(0) &= \tfrac{1}{2} \log [(1 + p_m{}^{-2\sigma})^2/(1 - p_m{}^{-2\sigma})], \end{aligned}$$

while

$$(61) \qquad 0 < a_m{}^\sigma(\omega) < a_m{}^\sigma(0) + a_m{}^\sigma(\tfrac{1}{2}\pi) \text{ for } 0 \leqq \omega < 2\pi,$$

since the two-fold symmetry of the convex curve $S_m{}^\sigma$ implies that

$$a_m{}^\sigma(\omega) = a_m{}^\sigma(\omega + \pi), \qquad a_m{}^\sigma(\omega) = a_m{}^\sigma(\pi - \omega).$$

Since $\sigma > 1$, it follows that the series

$$(62) \qquad r_k{}^\sigma(\omega) = \sum_{m=k+1}^{\infty} a_m{}^\sigma(\omega), \qquad (k = 0, 1, \cdots),$$

are uniformly convergent for $0 \leqq \omega < 2\pi$. It is known [31] that if σ and k are fixed, then the inequality $a_k{}^\sigma(\omega) > r_k{}^\sigma(\omega)$ holds for every ω if it holds for $\omega = \frac{1}{2}\pi$. Hence, if $\sigma^{(k)}$ denotes, for a fixed positive integer k, the obviously unique number such that

---

[30] Cf. Bohr and Courant [4] or Bohr and Jessen [6].

[31] Bohr and Jessen [6], p. 40.

$$\text{(63)} \qquad \arcsin p_k^{-\sigma} = \sum_{m=k+1}^{\infty} \arcsin p_m^{-\sigma} \text{ if } \sigma = \sigma^{(k)} > 1, \qquad (k = 1, 2, \cdots),$$

then it is seen from (60) that

$$\text{(64)} \qquad \sigma > \sigma^{(k)} \text{ if and only if } a_k^{\sigma}(\omega) > r_k^{\sigma}(\omega) \text{ for all } \omega,$$

where k is fixed.

The spectrum (56) of the infinite convolution (49) belonging to the curves (59) may be described as follows: The spectrum V^{σ} is [32] a closed bounded region bordered by a convex curve C^{σ} or a closed bounded ring-shaped region bordered by two disjoint convex curves C^{σ}, D^{σ} according as $1 < \sigma \leqq \sigma^{(1)}$ or $\sigma > \sigma^{(1)}$, where $\sigma^{(1)} = 1.764 \cdots$ is defined by (63). If $\sigma > \sigma^{(1)}$, i. e., if D^{σ} exists, let C^{σ} denote that of the two convex curves C^{σ}, D^{σ} which forms the outer boundary of V^{σ}. Let $\bar{\sigma}$ denote the obviously unique root of the equation

$$\text{(65)} \qquad p_1^{-\sigma} = \sum_{m=2}^{\infty} p_m^{-\sigma}$$

in the range $\sigma > 1$, so that [33] $\bar{\sigma} = 1.778 \cdots$, and so $\bar{\sigma} < \sigma^{(1)}$. Then, if $\sigma > \sigma^{(1)}$, i. e., if D^{σ} exists, the symmetric convex curve D^{σ} has [34] no corners or has corners on the x-axis and nowhere else according as $\sigma \geqq \bar{\sigma}$ or $\sigma < \bar{\sigma}$. Furthermore, if $\sigma \geqq \bar{\sigma}$, then the supporting function of D^{σ} is [35]

$$\text{(66)} \qquad D^{\sigma} : a_1^{\sigma}(\omega) - r_1^{\sigma}(\omega), \qquad (\sigma \geqq \sigma).$$

The supporting function of C^{σ} is [36] $r_0^{\sigma}(\omega)$ for every $\sigma > 1$.

Let

$$\text{(67)} \qquad e_1 = 1 \text{ and } e_m = \pm 1, \text{ where } m = 2, 3, \cdots,$$

so that there is a non-enumerable set of symbols $(e_1, e_2, \cdots)$. The relations (60) and (61) imply that the series

$$\text{(68)} \qquad r^{\sigma}(\omega; \{e_m\}) = \sum_{m=1}^{\infty} e_m a_m^{\sigma}(\omega), \qquad (\sigma > 1),$$

is, for every symbol

$$\text{(69)} \qquad (e_1, e_2, \cdots, e_m, \cdots) = (1, e_2, \cdots, e_m, \cdots),$$

---

[32] Bohr and Jessen [6].

[33] Bohr and Jessen [6], p. 43.

[34] Cf. Bohr and Jessen [6], p. 43 and Kershner [12].

[35] Kershner [11], where the result is proved for an arbitrary finite number of convex curves, but the proof holds for the infinite vectorial sum also.

[36] Haviland [8].

uniformly convergent for $0 \leqq \omega < 2\pi$. It is clear from (67), (68) and from the definition (62) of $r_1^{\sigma}(\omega)$ that $r^{\sigma}(\omega; \{e_m\})$ is the sum of $a_1^{\sigma}(\omega) - r_1^{\sigma}(\omega)$ and of a finite or infinite series the terms of which are some of the functions $a_m^{\sigma}(\omega) > 0$. This holds for every $\sigma > 1$. Now if $\sigma \geqq \bar{\sigma}\ (> 1)$, then $a_1^{\sigma}(\omega) - r_1^{\sigma}(\omega)$ is, by (66), the supporting function of the convex curve D^{σ}, while $a_m^{\sigma}(\omega)$ is the supporting function of the convex curve S_m^{σ} for any $\sigma > 1$. Hence the function (68) of ω is, by a classsical theorem,[37] the supporting function of a convex curve for every $\sigma \geqq \sigma$. This convex curve, which depends on the symbol (69), will be denoted by

$$C^{\sigma}(\{e_m\}), \qquad (\sigma \geqq \bar{\sigma}). \tag{70}$$

Now let,[38] for a fixed integer $l \geqq 2$,

$$\sigma > \text{Max}\ (\sigma^{(1)}, \cdots, \sigma^{(l)}) \tag{71}$$

and choose, for a fixed integer k which satisfies the inequalities $2 \leqq k \leqq l$, a pair of symbols (69) which will be denoted by

$$\{e_m\}_k^{\text{I}}, \quad \{e_m\}_k^{\text{II}}, \qquad (2 \leqq k \leqq l), \tag{72}$$

and have the property that the numbers $e_1 = 1, e_2, \cdots, e_{k-1}$ are respectively the same in both symbols, while $e_k = 1, e_{k+1} = -1, e_{k+2} = -1, e_{k+3} = -1, \cdots$ in $\{e_m\}_k^{\text{I}}$ and $e_k = -1, e_{k+1} = 1, e_{k+2} = 1, e_{k+3} = 1, \cdots$ in $\{e_m\}_k^{\text{II}}$. Thus there are 2^{k-2} pairs of symbols (72) for a fixed k, and so $2^{l-1} - 1$ pairs of symbols (72) for $2 \leqq k \leqq l$. Now if (72) is any of these $2^{l-1} - 1$ pairs of symbols and if (71) is satisfied, then the convex curve (70) which belongs to $\{e_m\} = \{e_m\}_k^{\text{II}}$ lies entirely within the convex curve (70) which belongs to $\{e_m\} = \{e_m\}_k^{\text{I}}$. In fact, it is clear from the definition of the convex curve (70) that the difference of the supporting functions of the convex curves

$$C^{\sigma}(\{e_m\}_k^{\text{I}}) \text{ and } C^{\sigma}(\{e_m\}_k^{\text{II}}) \tag{73a}$$

is $2a_k^{\sigma}(\omega) - 2r_k^{\sigma}(\omega)$, an expresssion which is positive for every ω in view of (64) and (71). This clearly [37] proves the statement. The *open* ring-shaped domain bordered by the two convex curves (73a) will be denoted by

$$H^{\sigma}(\{e_m\}_k), \qquad (2 \leqq k \leqq l). \tag{73b}$$

---

[37] Cf., e. g., Haviland [8].

[38] It will not be discussed which of the l absolute constants $\bar{\sigma}, \sigma^{(2)}, \cdots, \sigma^{(l)}$ is Max $(\bar{\sigma}, \sigma^{(1)}, \cdots, \sigma^{(l)})$ for a fixed l and what is the asymptotic behaviour of $\sigma^{(l)}$ and Max $(\bar{\sigma}, \sigma^{(1)}, \cdots, \sigma^{(l)})$ for large l.

It is seen by an obvious repetition of the above consideration that none of the infinitely many convex curves (70) has any point within the open set (73b). This implies that the $2^{l-1} - 1$ ring-shaped domains (73a) which belong to all pairs of symbols (72) are mutually disjoint.

Now it is easily seen that if $\delta^\sigma(x, y)$ denotes the density of the infinite convolution (49) in which $\phi_m{}^\sigma$ belongs to the curve (59), and if σ is chosen such that (71) is satisfied, then $\delta^\sigma(x, y)$ is a regular analytic function of the two real variables x, y within each of the $2^{l-1} - 1$ mutually disjoint ring-shaped open sets (73b).

First, let $\delta_2{}^\sigma(x, y)$ denote the density of the convolution $\psi_2{}^\sigma = \phi_1{}^\sigma * \phi_2{}^\sigma$ (cf. Section 3), where $\phi_1{}^\sigma$, $\phi_2{}^\sigma$ are the distribution functions belonging to the curves $S_1{}^\sigma$, $S_2{}^\sigma$ defined by (59). Since (71) implies that $\sigma \geqq \bar{\sigma}$, the difference $a_1{}^\sigma(\omega) - r_1{}^\sigma(\omega)$ is the supporting function of a convex curve; cf. (66). Since the function $r_2{}^\sigma(\omega)$ defined by (62) also is [37] the supporting function of a convex curve, it follows [37] that the same holds for the sum of $a_1{}^\sigma(\omega) - r_1{}^\sigma(\omega)$ and $r_2{}^\sigma(\omega)$. This means in view of (62) that

$$a_1{}^\sigma(\omega) - a_2{}^\sigma(\omega) \tag{74}$$

is the supporting function of a convex curve. It follows [39] that the vectorial sum $S_1{}^\sigma\ (+)\ S_2{}^\sigma$ is ring-shaped, with an inner boundary curve $D_2{}^\sigma$ which is free of corners and has the difference (74) as its supporting function. Hence it is seen from the last remark in Section 2 that the open set $R = R^\sigma$ on which the explicit representation (17) of $\delta_2 = \delta_2{}^\sigma$ is valid is identical with the open ring-shaped domain bordered by the two convex curves $C_2 = C_2{}^\sigma$, $D_2 = D_2{}^\sigma$, i. e., by the convex curves which have the supporting functions

$$a_1{}^\sigma(\omega) + a_2{}^\sigma(\omega)\,; \qquad a_1{}^\sigma(\omega) - a_2{}^\sigma(\omega). \tag{75}$$

This open ring-shaped domain will be denoted by $H_2{}^\sigma$, so that

$$H_2{}^\sigma = [S_1{}^\sigma\ (+)\ S_2{}^\sigma] - [C_2{}^\sigma + D_2{}^\sigma]. \tag{76}$$

Now the functions $\xi_m(\theta)$, $\eta_m(\theta)$ occurring in the definition (59) of $S_m{}^\sigma$ are regular analytic. Furthermore, the Jacobian (8) of the transformation (7), when expressed by means of (14) in terms of x, y, does not vanish on the set $H_2{}^\sigma$ (cf. the corresponding argument at the end of Section 6). Since the representation (17) of $\delta_2 = \delta_2{}^\sigma$ is valid on $H_2{}^\sigma$, it follows that $\delta_2{}^\sigma$ is a

[39] Kershner [11], Theorem III.

regular analytic function of the two real variables x, y on the open ring-shaped domain H_2^{σ}.

Let $\delta_n^{\sigma}(x, y)$ denote the density of the distribution function (50), so that, by (33) and (59),

$$(77)\quad \delta_n^{\sigma}(x,y) = \frac{1}{2\pi}\int_0^{2\pi} \delta_{n-1}^{\sigma}(x - g(\theta; p_n^{-\sigma}), y - h(\theta; p_n^{-\sigma}))\,d\theta, \text{ where } n \geqq 3.$$

It is obvious from the same consideration which has been applied before to H_2^{σ} that both functions

$$(78)\qquad a_1^{\sigma}(\omega) + a_2^{\sigma}(\omega) - \sum_{m=3}^{\infty} a_m^{\sigma}(\omega)\,; \qquad a_1^{\sigma}(\omega) - a_2^{\sigma}(\omega) + \sum_{m=3}^{\infty} a_m^{\sigma}(\omega)$$

are supporting functions of certain convex curves which form the outer and inner boundary of an open ring-shaped domain. On denoting this domain by H_n^{σ}, it is clear that H_{n+1}^{σ} is a subset of H_n^{σ} and that, for a fixed σ, the common part of all H_n^{σ} is the unique ring-shaped region (73b) which belongs to $k = 2$. Since the two functions (78) represent the supporting functions of the outer and inner boundaries of H_n^{σ}, it follows that if H_n^{σ} contains the point (x, y), then

$$(x - g(\theta; p_n^{-\sigma}), y - h(\theta; p_n^{-\sigma}))$$

is a point of H_{n-1}^{σ} for every θ. Hence the argument which led from (39) to (45_2) can be applied without any modification and shows that the density $\delta^{\sigma}(x, y)$ of (49) is, again in view of (32) and by virtue of Vitali's theorem, a regular analytic function of the two real variables x, y on the ring-shaped region (73b) which belongs to $k = 2$. Finally, the regular analyticity of $\delta^{\sigma}(x, y)$ in the remaining $2^{l-1} - 1 - 1$ open ring-shaped regions (73_2) follows in the same way; cf. the end of Section 5.

Appendix.[40] *On real almost periodic functions with linearly independent frequencies.* It is known [41] that if

$$(1)\qquad f(t) \sim \sum_{n=0}^{\infty} a_n \cos(\lambda_n t + \delta_n), \qquad (a_n > 0),$$

is almost periodic, then it has an asymptotic distribution function $\sigma(x)$, and that the spectrum of $\sigma(x)$ consists of the closure of the set of values attained

[40] The numbers of the formulae will refer to the formulae of this Appendix.

[41] Wintner [19], pp. 254-255 and p. 269; cf. also Wintner [20].

by $x = f(t)$ for $-\infty < t < +\infty$. In what follows, it will be assumed that the frequencies λ_k of (1) are linearly independent. Then the series

$$r = \sum_{n=0}^{\infty} a_n, \qquad (a_n > 0), \tag{2}$$

is convergent by a well-known theorem of Bohr, and the closure just mentioned is the interval $-r \leqq x \leqq r$. Furthermore, on placing

$$g(x) = \tfrac{1}{\pi}(1-x^2)^{-\frac{1}{2}}, \text{ if } |x| < 1; \quad g(x) = 0, \text{ if } |x| \geqq 1, \tag{3}$$

and defining a sequence $\sigma_0, \sigma_1, \cdots$ of distribution functions by means of the recursion formula

$$\sigma_n(x) = \int_{-\infty}^{+\infty} \sigma_{n-1}(x - a_n y)\, g(y)\, dy, \tag{4}$$

where

$$\sigma_0(x) = \int_{-\infty}^{x} g(y/a_0)\, dy/a_0, \tag{5}$$

it is known [42] that the asymptotic distribution function $\sigma(x)$ of (1) is represented for every x by

$$\sigma(x) = \lim_{n=\infty} \sigma_n(x), \tag{6}$$

and that [43] every derivative

$$\sigma^{(m)}(x) \text{ exists for } m = 1, 2, \cdots; \quad -\infty < x < +\infty. \tag{7}$$

Also, $\sigma_n(x)$ has [43] a continuous m-th derivative for every x, whenever

$$n \geqq 2m + 2. \tag{8}$$

Finally, as will be shown at the end of this paper,

$$\lim_{n=\infty} \sigma_n^{(m)}(x) = \sigma^{(m)}(x); \quad -\infty < x < +\infty, \qquad (m = 1, 2, \cdots). \tag{9}$$

Since the spectrum of $\sigma(x)$ is the interval $|x| \leqq r$, it is clear that $\sigma(x)$ cannot be regular analytic at $x = \pm r$. On the other hand, it is known [44]

---

[42] Wintner [16], pp. 312-316.
[43] Wintner [16], p. 315.
[44] Kershner and Wintner [14].

that if the coefficients $a_\varkappa$ of (1) satisfy certain inequalities, one can delimit within the spectrum $|x| \leqq r$ subregions of regular analyticity of $\sigma(x)$.

It was not proved so far that the end-points of the subregions determined in the paper just mentioned are singularities of $\sigma(x)$. There was not even known an example with any established singularity for $\sigma(x)$ in the interior of the spectrum. The object of the following considerations is to fill into these gaps. The method to be applied also gives some qualitative information as to the shape of the density curve $\sigma'(x)$; cf. (ii) and (iv) below. Needless to say, all singularities of $\sigma(x)$ are essential singularities in virtue of (7).

The results to be obtained may be summarized as follows:

(i) If

$$r < 2a_0, \tag{10}$$

then $\sigma(x)$ is regular analytic in the subinterval

$$-(2a_0 - r) < x < (2a_0 - r) \tag{11}$$

of the spectrum $|x| \leqq r$ and has at the end-points $x = \pm (2a_0 - r)$ of (11) essential singularities.

(ii) If (10) is satisfied, there exists for every $m \geqq 1$ an $\epsilon_m > 0$ such that the derivative $\sigma^{(m)}(x)$ is positive for $0 < x < 2a_0 - r + \epsilon_m$. On choosing $m = 2$ and $m = 3$, it follows that there exists an $\epsilon > 0$ such that the density $\sigma'(x)$ is monotone increasing and convex in the interval $0 \leqq x \leqq 2a_0 - r + \epsilon$.

(iii) If

$$r = 2a_0 \tag{12}$$

then $\sigma(x)$ has an essential singularity at $x = 0$.

(iv) If (12) is satisfied, there exists for every $m \geqq 1$ an $\epsilon_m > 0$ such that the derivative $\sigma^{(m)}(x)$ is positive for $0 < x < \epsilon_m$. On choosing $m = 2$ and $m = 3$, it follows that there exists an $\epsilon > 0$ such that the density $\sigma'(x)$ is monotone increasing and convex in the interval $0 \leqq x \leqq \epsilon$.

(It may be mentioned that on combining (ii) and (iv) with the representation of $\sigma(x)$ which is obtained [45] by Fourier inversion, there results the following curiosity: If (2) is a convergent series with positive terms which satisfy the inequality $r \leqq 2a_0$, then the number I_m defined by the definite Bessel integral

[45] Wintner [16], p. 315.

$$I_m = \int_0^\infty x^{2m} \prod_{n=0}^{\infty} J_0(a_n x)\, dx$$

is positive or negative according as the positive integer m is even or odd.)

First, it is easily seen from (3) and (4) that if $n \geqq 1$ and $p \geqq 0$, then, for every x,

$$\sigma_{p+n}(x) = \int_{-1}^{1} \cdots \int_{-1}^{1} \sigma_p(x - \sum_{i=1}^{n} a_{p+i} y_i)\, d\Omega, \tag{13}$$

where

$$d\Omega = \prod_{i=1}^{n} [\tfrac{1}{\pi} (1 - y_i^2)^{-\frac{1}{2}}\, dy_i]. \tag{14}$$

In particular,

$$\sigma_n(x) = \int_{-1}^{1} \cdots \int_{-1}^{1} \sigma_0(x - \sum_{i=1}^{n} a_i y_i)\, d\Omega. \tag{15}$$

It is understood that the integrals are n-fold, so that, from (14),

$$\int_{-1}^{1} \cdots \int_{-1}^{1} d\Omega = [\tfrac{1}{\pi} \int_{-1}^{1} (1 - y^2)^{-\frac{1}{2}}\, dy]^n = 1. \tag{16}$$

It is easily verified from (3), (5) and (15) that

$$\sigma_n(x) + \sigma_n(-x) = 1$$

for every n and x; hence

$$\sigma_n^{(m)}(x) = (-1)^{m+1} \sigma_n^{(m)}(-x), \qquad (n \geqq 0,\ m \geqq 1), \tag{17_n}$$

for all those x at which σ_n has an m-th derivative.

The proof of (i) proceeds as follows. Suppose that (10) is satisfied. Then it is clear from (2) that

$$-(a_0 - \sum_{i=1}^{n} a_i) < x < (a_0 - \sum_{i=1}^{n} a_i) \tag{18_n}$$

is an interval containing the interval (18_{n+1}) and the interval (11). It is also seen that if $p \geqq 0$ and $n \geqq 1$, and if x lies in the interval (18_{p+n}), then the number

$$x - \sum_{i=1}^{n} a_{p+i} y_i \tag{19}$$

lies in the interval (18_n) for all those $(y_1, \cdots, y_n)$ for which $|y_i| < 1$, where $i = 1, \cdots, n$. On combining this fact with (15), it is seen from the definition (5), (3) of $\sigma_0(x)$ that

(20) $\qquad \sigma_n(x)$ is regular analytic in (18_n),

and that

$$(21_n) \qquad \sigma_n^{(m)}(x) = \int_{-1}^{1} \cdots \int_{-1}^{1} \sigma_0^{(m)}(x - \sum_{i=1}^{n} a_i y_i)\, d\Omega \text{ in } (18_n).$$

Now (21_n) means that

$$(22_n) \qquad \sigma_{p+n}^{(m)}(x) = \int_{-1}^{1} \cdots \int_{-1}^{1} \sigma_p^{(m)}(x - \sum_{i=1}^{n} a_{p+i} y_i)\, d\Omega \text{ in } (18_{p+n})$$

holds for $p = 0$. On combining (21_p) with (21_{p+n}), it is seen from the remark made in connection with (19) that (22_n) holds for every p.

It is known [46] that

(23) $\qquad \sigma(x)$ is regular analytic in (11).

In order to discuss the Taylor series of $\sigma(x)$ in the vicinity of $x = 0$, notice first that all coefficients in the binomial expansion

$$(24) \qquad (1 - x^2)^{-\frac{1}{2}} = \sum_{n=0}^{\infty} b_n x^{2n}, \qquad |x| < 1,$$

are positive, so that, from (3) and (5),

$$(25) \qquad \sigma_0^{(m)}(x) > 0, \text{ if } 0 < x < a_0.$$

It is also seen that

$$(26) \qquad \sigma_0^{(2q-1)}(0) > 0, \qquad (q \geqq 1),$$

while, from (17_n),

$$(27) \qquad \sigma_n^{(2q)}(0) = 0, \qquad (n \geqq 0,\ q \geqq 1).$$

Now (21_n), (25) and (17_0) clearly imply that

$$(28_q) \qquad \sigma_n^{(2q-1)}(x) > 0 \text{ in } (18_n).$$

It is seen from (27) and (28_{q+1}) that the minimum of the function $\sigma_n^{(2q-1)}(x)$ in the interval (18_n) is attained at $x = 0$, so that, for every $q \geqq 1$,

[46] Kershner and Wintner [14].

(29_n) $$\sigma_n^{(2q-1)}(x) \geqq \sigma_n^{(2q-1)}(0) \text{ in } (18_n).$$

On placing in (22_n)

$$x = 0 \text{ and } m = 2q - 1,$$

it is seen from (29_p) and from the remark made in connection with (19) that

$$\sigma_{p+n}^{(2q-1)}(0) > \sigma_p^{(2q-1)}(0) \int_{-1}^{1} \cdots \int_{-1}^{1} d\Omega.$$

Hence, from (16) and (28_q),

(30) $$\sigma_{p+n}^{(2q-1)}(0) > \sigma_p^{(2q-1)}(0) > 0.$$

Now from (9) and (30)

(31) $$\sigma^{(2q-1)}(0) > \sigma_p^{(2q-1)}(0) > 0,$$

while from (9) and (27)

(32) $$\sigma^{(2q)}(0) = \sigma_p^{(2q)}(0) = 0,$$

so that

(33) $$\sigma^{(m)}(0) \geqq 0$$

for every m.

Let R denote the radius of convergence of the power series

(34) $$\sigma(x) = \sum_{m=0}^{\infty} \sigma^{(m)}(0) x^m/m!$$

and R_n that of the power series

(35) $$\sigma_n(x) = \sum_{m=0}^{\infty} \sigma_n^{(m)}(0) x^m/m!.$$

On writing n instead of p in (31) and (32), it is seen that

(36) $$R \leqq R_n$$

for every n. On the other hand, on keeping n fixed and choosing m in (21_n) sufficiently large, it is seen from (3) and (5) that the regular analytic function representing $\sigma_n(x)$ in the interval (18_n) is singular at the end-points

$$x = \pm \left(a_0 - \sum_{i=1}^{n} a_i\right)$$

of (18_n), so that

(37) $$R_n \leqq a_0 - \sum_{i=1}^{n} a_i.$$

On letting $n \to \infty$ in (36) and (37), it follows from (2) that

$$R \leqq 2a_0 - r; \text{ and } R > 0, \tag{38}$$

by (23). Now it is known that if the coefficient of a power series

$$f(z) = \sum_{m=0}^{\infty} c_m z^m \tag{39}$$

are real and non-negative, and if this power series is convergent for $|z| < s$ but not for $|z| > s$, then $z = s > 0$ is a singular point of the function $f(z)$ also when the series (39) is convergent at $z = s$ (Vivanti-Pringsheim). It follows, therefore, from (33), (23) and (38) that $R = 2a_0 - r$ and that the function (34) is singular at $x = 2a_0 - r$, and so, by (32), at $x = -(2a_0 - r)$ also. This completes the proof of (i).

Since (34) is valid for $0 < x < 2a_0 - r$, so is the expansion

$$\sigma^{(m)}(x) = \sum_{n=0}^{\infty} \sigma^{(n+m)}(0) x^n / n! . \tag{40}$$

On letting $x \to 2a_0 - r$ in (40), it is seen from (33) and (7) that, in view of Abel's continuity theorem, (40) is valid at $x = 2a_0 - r$ also. Consequently, $\sigma^{(m)}(x)$ is positive not only for $0 < x < 2a_0 - r$ but for $x = 2a_0 - r$ as well, and so, by (7), for $2a_0 - r < x < 2a_0 - r + \epsilon_m$ also, if $\epsilon_m > 0$ is sufficiently small. This proves (ii).

Now replace (10) by (12). It is clear that the proof of (31) and (32) is valid in the case (12) also, and that the radius of convergence R_n (> 0) of the power series (35) again satisfies the inequality (37). It is seen from (2), (12) and (37) that

$$R_n \to 0 \text{ as } n \to \infty. \tag{41}$$

On combining (31) and (32) with (41), it follows that the power series (34) is divergent for every $x \neq 0$. This proves (iii).

Finally, the proof of (iv) is the same as that of (ii).

The statement (9) used above may be proved as follows. For a fixed $m \geqq 1$, let $\rho_0(x)$ denote the convolution of the $2m + 5$ distribution functions

$$\sigma_0(xa_0/a_j), \text{ where } 0 \leqq j \leqq 2m + 4,$$

so that $\rho_0(x)$ has, in view of the remark made in connection with (8), a continuous $(m + 1)$-th derivative for every x. Furthermore, all derivatives of $\rho_0(x)$ vanish for sufficiently large $|x|$, since the spectrum of $\rho_0(x)$ is contained in the spectrum $|x| \leqq r$ of (6). Finally, on placing

$$\rho_n(x) = \sigma_0(xa_0/a_{n+2m+5}), \text{ where } n = 1, 2, \cdots,$$

the distribution function (6) is the infinite convolution of the distribution functions $\rho_0, \rho_1, \rho_2, \cdots$, since it is the infinite convolution of the distribution functions

$$\sigma_0(xa_0/a_k), \text{ where } k = 0, 1, 2, \cdots.$$

Hence (9) is implied by the following

LEMMA. If an infinite convolution of a sequence of distribution functions $\rho_0(x), \rho_1(x), \cdots$ converges to a distribution function $\sigma(x)$, and if $\rho_0(x)$ has for $-\infty < x < +\infty$ an absolutely integrable and bounded $(m+1)$-th derivative, then the convolution of the $n+1$ distribution functions $\rho_0(x), \cdots, \rho_n(x)$, has a continuous m-th derivative which tends, for $-\infty < x < +\infty$, to a continuous function representing the m-th derivative of $\sigma(x)$.

Now the proof of this Lemma requires but an obvious modification of the proof previously [47] given for the particular case $m = 1$.

THE JOHNS HOPKINS UNIVERSITY.

BIBLIOGRAPHY.

[1] H. Bohr, "Über die Funktion $\zeta'(s)/\zeta(s)$," *Journal für Mathematik*, vol. 141 (1912), pp. 217-234.

[2] ———, "Om Addition af uendelig mange konvekse Kurver," *Danske Videnskabernes Selskab*, Forhandlinger, 1913, pp. 325-366.

[3] ———, "Om Addition af konvekse Kurver med givne Sandsynlighedsfordelinger," *Matematisk Tidsskrift B*, 1923, pp. 10-15.

[4] ———, and R. Courant, "Neue Anwendungen der Theorie der diophantischen Approximationen auf die Riemannsche Zetafunktion," *Journal für Mathematik*, vol. 144 (1914), pp. 249-274.

[5] ———, and B. Jessen, "Om Sandsynlighedsfordelinger ved Addition af konvekse Kurver," *Danske Videnskabernes Selskab*, Skrifter (8), vol. 12, no. 3 (1929).

[6] ———, and B. Jessen, "On the distribution of the values of the Riemann zeta function," *American Journal of Mathematics*, vol. 58 (1936), pp. 35-44.

[7] C. Burrau, "Numerische Lösung der Gleichung . . . ," *Journal für Mathematik*, vol. 142 (1912), pp. 51-53.

[47] Wintner [18].

[8] E. K. Haviland, "On the addition of convex curves in Bohr's theory of Dirichlet series," *American Journal of Mathematics*, vol. 55 (1933), pp. 332-334.

[9] ———, and A. Wintner, "On the Fourier transforms of distributions on convex curves," *Duke Mathematical Journal*, vol. 2 (1936), pp. 712-721.

[10] B. Jessen and A. Wintner, "Distribution functions and the Riemann zeta function," *Transactions of the American Mathematical Society*, vol. 38 (1935), pp. 48-88.

[11] R. B. Kershner, "On the addition of convex curves," *American Journal of Mathematics*, vol. 58 (1936), pp. 737-746.

[12] ———, "On the values of the Riemann ζ-function on fixed lines $\sigma > 1$," *ibid.*, vol. 59 (1937), pp. 171-178.

[13] ———, and A. Wintner, "On symmetric Bernoulli convolutions," *ibid.*, vol. 57 (1935), pp. 541-548.

[14] ———, and A. Wintner, "On the asymptotic distribution of almost periodic functions with linearly independent frequencies," *ibid.*, vol. 58 (1936), pp. 91-94.

[15] ———, and A. Wintner, "On the boundary of the range of values of $\zeta(s)$," *ibid.*, vol. 58 (1936), pp. 421-425.

[16] A. Wintner, "Upon a statistical method in the theory of diophantine approximations," *ibid.*, vol. 55 (1933), pp. 309-331.

[17] ———, "On analytic convolutions of Bernoulli distributions," *ibid.*, vol. 56 (1934), pp. 659-663.

[18] ———, "On the differentiation of infinite convolutions," *ibid.*, vol. 57 (1935), pp. 363-366.

[19] ———, *Spektraltheorie der unendlichen Matrizen*, Leipzig, 1929.

[20] ———, "Über die Grundlagen des matrizenmechanischen Umdeutungsprinzips und über eine Spektraltheorie der Bohrschen fastperiodischen Funktionen." *Zeitschrift für Physik*, vol. 48 (1928), pp. 149-161.

ON THE ORDER OF THE COEFFICIENTS OF A UNIVALENT FUNCTION.[1]

By M. S. ROBERTSON.

Introduction. If $f(z) = z + \sum_2^\infty a_n z^n$ is holomorphic and univalent for $|z| < 1$ then it is known [2] that $a_n = O(n)$. It has been conjectured [3] that $|a_n| \leqq n$. For $n = 2$ and 3 this conjecture is known to be true, equality occurring only for the function $z(1 + ze^{i\alpha})^{-2}$, α real. It is the purpose of the present paper to show that the univalent functions $f(z)$ for which

$$\limsup \frac{\log |a_n|}{\log n} = 1 \tag{1}$$

is true are also closely related to this same extremal function $z(1 + ze^{i\alpha})^{-2}$.

We show first of all that the only univalent function $f(z)$ for which

$$\left[f(x) - f\left(\frac{x - z}{1 - \bar{x}z}\right)\right] \cdot [(1 - |x|^2)f'(x)]^{-1} \equiv z(1 + ze^{i\alpha})^{-2} \tag{2}$$

for any x of modulus less than unity, and α a real number, is true is the very special function

$$f(z) \equiv \frac{z + (\bar{x} - xe^{2i\alpha})(\bar{x}e^{2i\alpha} - 1)^{-1}z^2}{\{1 - (\bar{x} + e^{i\alpha})(xe^{i\alpha} + 1)^{-1}z\}^2}. \tag{3}$$

Secondly, we show that a necessary condition that $f(z)$ satisfy (1) is that $f(z)$ be a rather special type of function satisfying (2) in an asymptotic sense. More precisely, in order that $f(z)$ should satisfy (1) it must necessarily possess the following very special property: there shall exist a sequence of points x_i inside the unit circle tending to a singularity of $f(z)$ on the circumference, and a real number α, so that

$$\lim_{|x_i| \to 1} \left[f(x_i) - f\left(\frac{x_i - z}{1 - \bar{x}_i z}\right)\right] \cdot [(1 - |x_i|^2)f'(x_i)]^{-1} \equiv z(1 + ze^{i\alpha})^{-2}. \tag{4}$$

---

[1] Received September 16, 1936; revised October 26, 1936.

[2] See J. E. Littlewood, *Proceedings of the London Mathematical Society* (2), vol 23 (1925), p. 482.

[3] See L. Bieberbach, *Sitzber. kgl. Akad. Berlin*, 1916, s. 940-955.

The following lemma is due to L. Bieberbach.[4]

LEMMA 1. *If* $f(z) = z + \sum_2^\infty a_n z^n$ *is holomorphic and univalent for* $|z| < 1$ *then* $|a_2| \leqq 2$ *where equality holds for, and only for, the function* $z(1 + ze^{i\alpha})^{-2}$, α *real.*

LEMMA 2. *Given* $f(z) = z + \sum_2^\infty a_n z^n$ *holomorphic for* $|z| < 1$, *and* x_0 *any fixed complex number such that* $|x_0| < 1$. *If we have identically, for some real constant* α,

$$(5) \qquad \left[f(x_0) - f\left(\frac{x_0 - z}{1 - \bar{x}_0 z}\right)\right] \cdot [(1 - |x_0|^2) \cdot f'(x_0)]^{-1} \equiv z(1 + e^{i\alpha}z)^{-2}$$

then $f(z)$ *must be identically equal to the function*

$$(6) \qquad f_0(z) \equiv \frac{z + (\bar{x}_0 - x_0 e^{2i\alpha})(x_0^2 e^{2i\alpha} - 1)^{-1} z^2}{\{1 - (\bar{x}_0 + e^{i\alpha})(x_0 e^{i\alpha} + 1)^{-1} z\}^2}.$$

Proof. Replacing z by $(x_0 - z)/(1 - \bar{x}_0 z)$ in (5) we obtain

$$(7) \qquad \frac{f(x_0) - f(z)}{(1 - |x_0|^2) f'(x_0)} \equiv (x_0 - z)(1 - \bar{x}_0 z) \cdot \{1 + x_0 e^{i\alpha} - (\bar{x}_0 + e^{i\alpha})z\}^{-2}.$$

Since $f(0) = 0$, $f'(0) = 1$ we have

$$(8) \qquad f(x_0)[(1 - |x_0|^2) f'(x_0)]^{-1} = x_0(1 + x_0 e^{i\alpha})^{-2}$$

and

$$(9) \qquad [(1 - |x_0|^2) f'(x_0)]^{-1} = (1 - |x_0|^2)(1 - x_0 e^{i\alpha})(1 + x_0 e^{i\alpha})^{-3}.$$

Making the substitutions (8) and (9) in (7) we obtain for $f(z)$ the function $f_0(z)$ in (6) as required. We remark in passing that $f_0(z)$ is necessarily univalent for $|z| < 1$ since the right-hand side of the identity (5) is a function univalent in the unit circle, and since $(x_0 - z)/(1 - \bar{x}_0 z)$ is linear in z, mapping the unit circle onto itself.

LEMMA 3. *The function* $f_0(z)$ *of* (6) *above is such that it possesses a sequence of points* x_i *tending along a radius of the unit circle to the singularity of* $f_0(z)$ *on* $|z| = 1$ *for which* (4) *holds.*

This lemma may be easily verified by the reader.

THEOREM 1. *Let* $f(z) = z + \sum_2^\infty a_n z^n$ *be holomorphic and univalent for*

[4] See L. Bieberbach, *loc. cit.*

$|z|<1$. *Let E denote the set of points on $|z|=1$ at which $f(z)$ fails to be regular. Suppose that $f(z)$ has the property that there exists no sequence of points x_i of modulus inferior to unity, having a limit point in E for which* (4) *holds. Then there exists a $\mu>0$, depending upon $f(z)$ but not upon r, and a constant $A(\mu)$ independent of r such that*

$$|f(re^{i\theta})| \leqq \frac{A(\mu)\cdot r}{(1-r)^{2-\mu}}, \qquad \mu<2.$$

Proof. Let x be any point inside the unit circle. Let

$$F(z;x) = \left[f(x) - f\left(\frac{x-z}{1-\bar{x}z}\right)\right]\cdot[(1-|x|^2)f'(x)]^{-1}$$
$$= z + b_2(x)z^2 + \cdots.$$

Since $f(z)$ is univalent and holomorphic for $|z|<1$ so also is $F(z;x)$. On account of the hypothesis of our theorem relating to (4) we must have $f(z) \not\equiv f_0(z)$ by Lemma 3. Hence $F(z;x) \not\equiv z(1+ze^{i\alpha})^{-2}$ for any real value of α by Lemma 2. It follows from Lemma 1 that $|b_2(x)|<2$. We may then write

$$\underset{|x|<1}{l.\,u.\,b.}\,|b_2(x)| = 2-\mu(f), \qquad 0 \leqq \mu(f) \leqq 2.$$

We wish now to show that $\mu(f)>0$. Suppose, if possible, $\mu(f)=0$. Then there exists a sequence of points $\{x_i\}$ of modulus inferior to unity, converging to a point x_0, $|x_0| \leqq 1$, and a corresponding sequence of functions

$$F_i(z) \equiv F(z;x_i) = z + b_2(x_i)z^2 + \cdots$$

where $\lim_{i\to\infty}|b_2(x_i)| = 2$. Since the functions $F(z;x_i)$ are univalent for $|z|<1$ and therefore form a normal family in any domain interior to the unit circle, there exists an analytic function

$$F_0(z) = z + b_2{}^0z^2 + \cdots$$

such that the functions $F(z;x_i)$ converge uniformly to $F_0(z)$ in any domain completely interior to the unit circle. Moreover,

$$|b_2{}^0| = \lim_{x_i\to x_0}|b_2(x_i)| = 2.$$

Since the convergence of the functions is uniform and $F_0(z)$ is not a constant,[5] $F_0(z)$ is univalent for $|z|<1$. By Lemma 1

[5] See E. C. Titchmarsh, *On the Theory of Functions*, Oxford, 1932, page 200, 6.44.

$$F_0(z) \equiv z(1 + ze^{i\alpha})^{-2}$$

for some real value α, since $| b_2^0 | = 2$. There are now several cases to consider, depending upon whether $| x_0 | < 1$ or $| x_0 | = 1$. Suppose first $| x_0 | < 1$. Then

$$\frac{f(x_0) - f\left(\dfrac{x_0 - z}{1 - \bar{x}_0 z}\right)}{(1 - | x_0 |^2) f'(x_0)} \equiv \lim_{x_i \to x_0} F(z; x_i) \equiv F_0(z) \equiv z(1 + ze^{i\alpha})^{-2}.$$

But this is impossible unless $f(z)$ be the function $f_0(z)$ of (6) which by hypothesis we have excluded (since it follows from (4) and Lemma 3 as we remarked above).

Secondly, suppose $| x_0 | = 1$ and that $x_0 \subset E$. This case is clearly impossible by the hypothesis of our theorem. Thirdly, suppose $| x_0 | = 1$ and x_0 is a point of regularity of $f(z)$. Since

$$\frac{x_i - z}{1 - \bar{x}_i z} = x_i + \frac{(| x_i |^2 - 1)z}{1 - \bar{x}_i z},$$

$$F_i(z) = \frac{z}{1 - \bar{x}_i z} + \sum_{n=2}^{\infty} \frac{f^{(n)}(x_i)}{n!\, f'(x_i)} \cdot (| x_i |^2 - 1)^{n-1} \left(\frac{z}{1 - \bar{x}_i z}\right)^n.$$

Since for every i both x_0 and x_i are points of regularity for $f(z)$ there exists a positive ρ independent of i such that $f(z)$ is regular in each circle of radius ρ having a centre at x_0 or x_i. Hence by Cauchy's inequality we have for some constant M depending only upon $f(z)$ and not upon n or i

$$\left| \frac{f^{(n)}(x_i)}{n!} \right| \leqq \frac{M}{\rho^n},$$

$$\therefore \lim_{i \to \infty} \left| F_i(z) - \frac{z}{1 - \bar{x}_0 z} \right| \leqq \lim_{i \to \infty} (1 - | x_i |^2) P(z; x_i) = 0,$$

where $P(z; x_i)$ is a bounded quantity for all i. Hence in case x_0 is a point of regularity of $f(z)$ on $| z | = 1$ we have

$$F_0(z) \equiv \lim_{i \to \infty} F(z; x_i) \equiv \frac{z}{1 - \bar{x}_0 z}$$

which contradicts our assumption that $F_0(z) \equiv z(1 + ze^{i\alpha})^{-2}$. Consequently we have shown under the assumptions of our theorem that $\mu(f) > 0$.

We now can write

$$| b_2(x) | = \left| \tfrac{1}{2} \frac{f''(x)}{f'(x)} (1 - | x |^2) - \bar{x} \right| \leqq 2 - \mu.$$

Let $x = re^{i\theta}$. On integrating we obtain

$$|\log f'(x) + \log(1-r^2)| \leqq \log\left(\frac{1+r}{1-r}\right)^{2-\mu}$$

$$|f'(re^{i\theta})| \leqq \frac{(1+r)^{1-\mu}}{(1-r)^{3-\mu}}$$

$$|f(re^{i\theta})| \leqq \int_0^r \frac{(1+r)^{1-\mu}}{(1-r)^{3-\mu}}\,dr < 2\int_0^r \frac{dr}{(1-r)^{3-\mu}}$$

If $\mu \neq 2$ then

$$|f(re^{i\theta})| \leqq \frac{2}{2-\mu}\left[\frac{1-(1-r)^{2-\mu}}{(1-r)^{2-\mu}}\right] < \frac{A(\mu)\cdot r}{(1-r)^{2-\mu}}.$$

If $\mu = 2$ then

$$|f(re^{i\theta})| \leqq \tfrac{1}{2}\log\left(\frac{1+r}{1-r}\right).$$

This completes the proof of our theorem.

THEOREM 2. *Let* $f(z) = z + \sum_2^\infty a_n z^n$ *satisfy the hypotheses of Theorem* 1. *Then there exist two positive constants* δ *and* $A(\delta)$ *depending upon* $f(z)$, *but not upon* r, *where* $0 < \delta < 1$ *such that*

$$\frac{1}{2\pi}\int_0^{2\pi} |f(re^{i\theta})|\,d\theta < \frac{A(\delta)\cdot r}{(1-r)^{1-\delta}}.$$

Proof. By Theorem 1 we can find a value of δ for which $0 < \delta < 1$ so that

$$|f(re^{i\theta})| < \frac{A(\delta)\cdot r}{(1-r)^{2-\delta}}$$

where $A(\delta)$ is a constant independent of r. If we define

$$F(\zeta) = \sqrt{f(\zeta^2)}, \qquad |\zeta| = \rho = r^{1/2},$$

then

$$|F(\zeta)| \leqq A(\delta)\cdot\rho(1-\rho^2)^{-1+\delta/2}.$$

Further, we have the following inequality due to J. E. Littlewood.[6]

$$\frac{1}{2\pi}\int_0^{2\pi} |f(re^{i\theta})|\,d\theta \leqq 2\int_0^\rho \max_{|\zeta|=\rho} |F(\zeta)|^2 \frac{d\rho}{\rho}.$$

$$\therefore \quad \frac{1}{2\pi}\int_0^{2\pi} |f(re^{i\theta})|\,d\theta \leqq A(\delta)\int_0^\rho \frac{2\rho d\rho}{(1-\rho^2)^{2-\delta}}, \qquad \rho^2 = r$$

$$< \frac{A(\delta)\cdot r}{(1-r)^{1-\delta}}$$

where $A(\delta)$ is not necessarily the same constant in each place it appears above.

[6] See J. E. Littlewood, *loc. cit.*

THEOREM 3. *Let* $f(z) = z + \sum_2^\infty a_n z^n$ *satisfy the hypothesis of Theorem* 1. *Then there exists a* $\delta = \delta(f) > 0$, *depending upon* $f(z)$ *but not upon* n *such that*

$$a_n = O(n^{1-\delta}).$$

Proof. Since

$$|a_n| r^n \leqq \frac{1}{2\pi} \int_0^{2\pi} |f(re^{i\theta})| \, d\theta < \frac{A(\delta) r}{(1-r)^{1-\delta}}$$

we obtain for $r = 1 - 1/n$, $n > 1$,

$$\begin{aligned} |a_n| &\leqq A(\delta) \left(1 - \frac{1}{n}\right)^{-n+1} \cdot n^{1-\delta} \\ &\leqq A(\delta) \left(1 + \frac{1}{n-1}\right)^{n-1} \cdot n^{1-\delta} \\ &< eA(\delta) \cdot n^{1-\delta}, \qquad n > 1. \end{aligned}$$

We now have immediately

COROLLARY. *If* $f(z)$ *satisfies the condition* (1) *it must then also possess the property* (4).

YALE UNIVERSITY.

INDEPENDENT SETS OF POSTULATES FOR ABELIAN GROUPS AND FIELDS IN TERMS OF THE INVERSE OPERATIONS.[1]

By David G. Rabinow.

In a previous paper [2] I discussed an independent set of postulates for a group in terms of the inverse operation. In part I of the present paper I shall consider a simplified set of independent postulates for an Abelian group in terms of the inverse operation, while in part II independent postulates for a field in terms of the inverse operations will be developed.

For convenience let us here recapitulate the results of the previous paper referred to. The base is $(K, -)$ where K is a class of elements $a, b, c, \cdots$ and $-$ is a binary operation.

Postulate I. a in K and b in K imply $a - b$ in K.
Postulate II. $a - a = b - b$.
Definition I. $z = a - a$.
Definition II. $a' = z - a$.
Postulate III. $(a - b') - c' = a - (b - c')'$.
Postulate IV. $a'' = a$.
Postulate V. $a - b' = b - a'$.
Definition III. $a + b = a - b'$.

The results obtained were:

Result I. Any system $(K, -)$ which satisfies Postulates I through IV is a group with respect to $+$ as defined in Definition III.

Result II. Any system $(K, -)$ which satisfies Postulates I through V is an Abelian group with respect to the operation $+$.

Result III. An Abelian group is also given by I, II, III, V, and 11 where Theorem 11 is: $a - z = a$.

Result IV. Theorem 11 is deducible from I, II, III, and IV.

Result V. I, II, III, IV, and V are independent. Also I, II, III, V, and 11 are independent.

Hereafter we shall refer to these results as RI, RII, RIII, RIV, and RV respectively.

[1] Received June 16, 1936; revised October 5, 1936.

[2] "Independent set of postulates for a group in terms of the inverse operation," offered to the *Bulletin of the American Mathematical Society*.

PART I.

It is obvious from the above set of postulates that the definition of an Abelian group in terms of the inverse operation is much more complicated than the definition in terms of the direct operation. I shall now consider an independent set of postulates for an Abelian group in terms of the inverse operation which is as simple as the customary definition [3] in terms of the direct operation.

Let us consider the following set of postulates in connection with the base $(K, —)$ where K is a class of elements $a, b, c, \cdots$ and $—$ is a binary operation.

Postulate 1. a in K and b in K imply $a — b$ in K.

Postulate 2. $(a — b) — c = (a — c) — b$.

Postulate 3. $a — (a — b) = b$.

The following theorems are deducible from Postulates 1, 2 and 3.

THEOREM 1. *If* $a — b = a — c$, *then* $b = c$.

by hypothesis $a — b = a — c$

by Postulate 1 $a — (a — b) = a — (a — c)$

hence by Postulate 3 $b = c$.

THEOREM 2. $b — b = c — c$ *for every element.*

by Postulate 3 $b = a — (a — b)$

by Postulate 1 $b — b = [a — (a — b)] — b$

by Postulate 2 $= (a — b) — (a — b)$

hence by Postulate 1 $(a — b) — (b — b) = (a — b) — [(a — b) — (a — b)]$

by Postulate 3 $= a — b$ (I)

now by Postulate 2 $(a — a) — b = (a — b) — a$

hence $(a — b) — [(a — a) — b] = (a — b) — [(a — b) — a]$

by Postulate 3 $= a$ (II)

hence by (I) $[(a — b) — (b — b)] — [(a — a) — b] = a$

by Postulate 2 $\{(a — b) — [(a — a) — b]\} — (b — b) = a$

hence by (II) $a — (b — b) = a$

[3] For bibliographical references the reader may consult: E. V. Huntington, "Note on the definition of abstract groups and fields by sets of independent postulates," *Transactions of the American Mathematical Society*, vol. 6 (1905), pp. 181-193; and R. Garver, "Postulates for special types of groups," *Bulletin of the American Mathematical Society*, vol. 42 (1936), pp. 125-129.

however b is any element of K, therefore

$$a - (c - c) = a, \text{ where } c \text{ is any element of } K$$

hence $a - (b - b) = a - (c - c)$

hence by Theorem 1 $b - b = c - c.$

DEFINITION 1. $z = a - a.$

DEFINITION 2. $a' = z - a.$

THEOREM 3. $a'' = a.$

by Postulate 3 $b - (b - a) = a$ for all elements b

by Definition 1 $z - (z - a) = a$

by Definition 2 $a'' = a.$

THEOREM 4. $a - b = (b - a)'.$

by Postulate 3 $b - (b - a) = a$

by Postulate 1 $[b - (b - a)] - b = a - b$

by Postulate 2 $(b - b) - (b - a) = a - b$

by Definition 1 $z - (b - a) = a - b$

by Definition 2 $(b - a)' = a - b.$

THEOREM 5. $a - b = b' - a'.$

by Definition 2 $b' - a' = (z - b) - (z - a)$

by Postulate 2 $= [z - (z - a)] - b$

by Theorem 3 $= a - b.$

THEOREM 6. $(a - b') - c' = a - (b - c)'.$

by Theorem 4 $(b - c)' = c' - b$

by Postulate 1 $a - (b - c)' = a - (c' - b)$

by Theorem 4 $= [(c' - b) - a]'$

by Postulate 2 $= [(c' - a) - b]'$

by Theorem 4 $= b - (c' - a)$

by Theorem 4 $= b - (a - c)'$

by Theorem 5 $= (a - c') - b'$

by Postulate 2 $= (a - b') - c'.$

THEOREM 7. $a - b' = b - a'.$

by Theorem 6 $(a - b') - (b - a') = a - [b - (b - a')]'$

by Postulate 3 $= a - a''$

by Theorem 3 $= a - a$

by Definition 1 $= z$

now by Definition 1 $(a - b') - (a - b') = z$
hence $(a - b') - (a - b') = (a - b') - (b - a')$
by Theorem 1 $a - b' = b - a'$.

DEFINITION 3. $a + b = a - b'$.

However by RII we know that Postulate 1, Theorems 2, 3, 6, and 7 are the postulates for an Abelian group with respect to the operation $+$ defined in Definition 3. Hence Postulates 1, 2, and 3 define an Abelian group with respect to $+$. It remains to show that Postulates 1, 2, and 3 are deducible from RII; that is, we have shown that any system $(K, -)$ which satisfies Postulates 1, 2, and 3 is an Abelian group and we must show conversely, that any Abelian group satisfies Postulates 1, 2, and 3.

THEOREM 8. a in K and b in K imply $a - b$ in K.
by Postulate I this is true.

THEOREM 9. $(a - b) - c = (a - c) - b$.
by Postulate III $(a - b) - c = a - (b' - c)'$
by Postulate V $= a - (c' - b)'$
by Postulate III $= (a - c) - b$.

THEOREM 10. $a - (a - b) = b$.
by Theorem 8 $a - (a - b)$ is in K and hence $a - (a - b) = X$
by Theorem 8 $[a - (a - b)] - (a - b)' = X - (a - b)'$
by Postulate III $a - [(a - b)' - (a - b)']' = X - (a - b)'$
by Definitions I and II $a - z = X - (a - b)'$
by RIV and Postulate III $a = (X - a') - b$
by Theorem 8 $a - b' = [(X - a') - b] - b'$
hence by Postulate III, Definitions I and II, and RIV
$a - b' = X - a'$
by Postulate V $b - a' = X - a'$
by Theorem 8 $(b - a') - a = (X - a') - a$
by Postulate III, Definitions I and II, and RIV
$b = X$.

The postulates are examined for independence by the usual method of exhibiting examples of systems $(K, -)$ which fail to satisfy the correspondingly numbered postulate but satisfy the remaining postulates.

Example 1) K is the class of all prime positive integers with $a - b = a/b$.
Example 2) K is the class of all positive integers with $a - b = b$.
Example 3) K is the class of all positive integers with $a - b = a + b$.

PART II.

Let us consider the following set of postulates in connection with the base $(K, —, o)$ where K is a class of elements $a, b, c, \cdots$ and $—, o$ are binary operations.

Postulate E. K contains at least two distinct elements.

Postulate 11. a in K and b in K imply $a — b$ in K.

Postulate 12. $a — a = b — b$.

Definition 1. $z = a — a$.

Definition 2. $a' = z — a$.

Postulate 13. $(a — b') — c' = a — (b — c')'$.

Postulate 14. $a'' = a$.

Postulate 15. $a — b' = b — a'$.

Definition 3. $a + b = a — b'$.

Postulate 16. a in K and b in K and $b \neq z$ imply aob in K.

Postulate 17. $aoa = bob$ if aoa and bob are in K.

Definition 4. $U = aoa$ if aoa is in K.

Definition 5. $a^* = Uoa$ if Uoa is in K.

Postulate 18. 1) If U exists then $U \neq z$.

2) If a^* exists then $a^* \neq z$ (provided $a \neq z$).

Postulate 19. If $a, b, c, b^*, c^*, aob^*, (boc^*)^*, (aob^*)oc^*$ and $ao(boc^*)^*$ are in K, then $(aob^*)oc^* = ao(boc^*)^*$.

Postulate 20. If a, a^* and a^{**} are in K, then $a^{**} = a$.

Postulate 21. If a, b, a^*, b^*, aob^* and boa^* are in K, then $aob^* = boa^*$.

Postulate 22. 1) If $a, b, c, b — c', b^*, c^*, (b — c')^*, aob^*, aoc^*, ao(b — c')^*$ and $aob^* — (aoc^*)'$ are in K, then $ao(b — c')^* = aob^* — (aoc^*)'$.

2) If $a, b, c, b — c', a^*, boa^*, (coa^*)', (b — c')oa^*$ and $boa^* — (coa^*)'$ are in K, then $(b — c')oa^* = boa^* — (coa^*)'$.

Definition 6. 1) $a \cdot b = aob^*$ if $b \neq z$.

2) $a \cdot z = z$.

Note 1. Postulates 15, 20 and 22.1 are redundant. They have been retained to keep the list of postulates symmetrical.

LEMMA. *The element U must exist.*

For by Postulate E and Definition 1 there exists an element distinct from z. Let this element be a. Then by Postulate 16 aoa exists, and hence by Definition 4 the element U exists.

THEOREM 23. $(K, -)$ forms an Abelian group with respect to the *operation* $+$ *defined in Definition* 3. *This follows from* RII.

THEOREM 24. *a in K and b in K imply* $a \cdot b$ *in K.*

Case I. $b = z$.

by Definition 6.2	$a \cdot z = z$ for all a.

Case II. $b \neq z$.

by hypothesis	$b \neq z$
by Postulate 18.2	$b^* \neq z$
by Postulate 16	b^* is in K
by Postulate 16	aob^* is in K
by Definition 6.1	$a \cdot b$ in K.

THEOREM 25. *If* $a - b = a - c$ *then* $b = c$.

by hypothesis	$a - b = a - c$
by Definition 2	$(a - b)' = (a - c)'$
by Postulate 11	$a' - (a - b)' = a' - (a - c)'$
by Postulate 13	$(a' - a') - b = (a' - a') - c$
by Definition 1	$z - b = z - c$
by Postulate 14	$b = c$.

THEOREM 26. $z \cdot a = z$.

Case I. $a = z$.

by Definition 6.2	$a \cdot z = z$ for all a
hence	$z \cdot z = z$.

Case II. $a \neq z$.

by Definition 6.1	$a \cdot b = aob^*$ if $b \neq z$
hence	$z \cdot a = zoa^*$
by Theorem 23	$= zoa^* - z$
now by Definitions 1 and 2	$zoa^* = (z - z)oa^* = (z - z')oa^*$
by Postulate 22.2	$= zoa^* - (zoa^*)'$
hence	$zoa^* - z = zoa^* - (zoa^*)'$
by Theorem 25	$(zoa^*)' = z$

hence by Postulate 14 and Definition 2 $zoa^* = z$.

THEOREM 27. $a \cdot b = b \cdot a$.

Case I. $a = b = z$.

by Theorem 26	$z \cdot z = z \cdot z = z$

Case II. $b = z$, $a \neq z$.

by Definition 6.2 $a \cdot z = z$

by Theorem 26 $z \cdot a = z$.

Case III. $a \neq z$, $b \neq z$.

by Postulate 18.2 $a^* \neq z$, $b^* \neq z$

by Postulate 16 aob^* and boa^* are in K

by Postulate 21 $aob^* = boa^*$

by Definition 6.1 $a \cdot b = b \cdot a$.

THEOREM 28. $U = U^*$.

by Postulate 18.1 $U \neq z$

by Definition 5 $U^* = UoU$

by Definition 4 $= U$.

THEOREM 29. $boU = b$.

Case I. $b = z$.

Since by Postulate 18.1 $U \neq z$ then boU is in K by Postulate 16.

by Theorem 28 $boU = boU^*$

hence $zoU = zoU^*$

by Theorem 26 $= z$.

Case II. $b \neq z$.

by Theorem 28 $boU = boU^*$

by Definition 4 $= bo(b^*ob^*)^*$ (A)

now since by hypothesis $b \neq z$ hence by Postulate 18.2 $b^* \neq z$ and $b^{**} \neq z$

by Postulate 19, (A) $boU = (bob^{**})ob^*$

by Postulate 20 $= (bob)ob^*$

$= Uob^* = b^{**} = b$ by Definitions 4 and 5 and Postulate 20.

THEOREM 30. $a \neq z$ and $b \neq z$ *imply* $aob \neq z$.

by Postulate 20 $aob = aob^{**}$

now suppose $aob = z$ then

by Theorem 26 $(aob^{**})ob^* = zob^* = z$

$= ao(b^*ob^*)^* = aoU^* = aoU = a$ by Postulate 19, Definition 4, and Theorem 29.

hence $a = z$

but this contradicts the hypothesis that $a \neq z$,

hence $aob \neq z$.

THEOREM 31. $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.

Case I. $z = c$, $a \neq z$, $b \neq z$.

by Theorem 24	ab is in K
by Definition 6.2	$(a \cdot b) \cdot z = z$
by Definition 6.2	$b \cdot z = z$
by Definition 6.2	$a \cdot z = z$.

Case II. $b = z$.

by Definition 6.2	$az = z$
by Theorem 26	$z \cdot c = z$
by Theorem 26	$z \cdot c = z$
by Definition 6.2	$a \cdot z = z$.

Case III. $a = z$.

by Theorem 26	$z \cdot b = z$
by Theorem 26	$z \cdot c = z$
by Theorem 24	$b \cdot c$ is in K
by Theorem 26	$z \cdot (b \cdot c) = z$.

Case IV. $a = b = z$ and other combinations follow from Definition 6.2 and Theorem 26.

Case V. $a \neq z$, $b \neq z$, $c \neq z$.

by hypothesis	$a \neq z$, $b \neq z$, $c \neq z$
hence by Postulate 18.2 and Theorem 30	$a^* \neq z$, $b^* \neq z$, $c^* \neq z$, $boc^* \neq z$
by Postulate 19	$ao(boc^*)^* = (aob^*)oc^*$
by Definition 6.1	$a \cdot (b \cdot c) = (a \cdot b) \cdot c$.

THEOREM 32.1. *There exists an element* $u \neq z$ *such that* $a \cdot u = a$.

by Theorem 29	$aoU = a$ for all a
by Theorem 28	$aoU = aoU^*$
by Postulates 18.1 and 18.2	$U^* \neq z$
by Definition 6.1	$a \cdot U = a$ for all a.

THEOREM 32.2. $u \cdot a = a$ *where* u *is the* u *of Theorem* 32.1.

Case I. $a = z$.

by Definition 6.2	$U \cdot z = z$.

Case II. $a \neq z$.

by Definition 6.1	$U \cdot a = Uoa^*$
	$= a^{**} = a$ by Definition 5 and Postulate 20.

THEOREM 33. *If $a \neq z$ then there exists an element x such that $ax = b$.*

Case I. $b = z$.

by Definition 6.2 $a \cdot z = z$.

Case II. $b \neq z$.

by Postulate 20 $b = U o b^*$

by Definition 4 $= (aoa) o b^*$

by Postulate 19 $= ao(a^* o b^*)^*$

by hypothesis and Theorem 30 $(a^* o b^*)^* \neq z$

hence by Definition 6.1 $b = a \cdot x$ where $x = a^* o b^*$.

THEOREM 34. *If $a \neq z$ then there exists an element y such that $ya = b$.*

Case I. $b = z$.

by Theorem 26 $z \cdot a = z$.

Case II. $b \neq z$.

by Theorem 29 $b = bo U$

$= bo(a^* o a^*)^*$ by Definition 4 and Theorem 28.

by hypothesis $a \neq z$, hence by Postulate 18.2 $a^* \neq z$, then

by Postulate 19 $bo U = (boa) o a^*$

by Definition 6.1 $b = y \cdot a$ where $y = boa$ (A).

THEOREM 35. *If $y_1 \cdot a = b$ and $y_2 \cdot a = b$, then $y_1 = y_2$ if $a \neq z$.*

Case I. $b = z$.

by hypothesis $y_1 \cdot a = y_2 \cdot a$. Now if $y_1 \neq z$ or $y_2 \neq z$ and by hypothesis $a \neq z$, then by Theorem 30 either $y_1 o a^* \neq z$ or $y_2 o a^* \neq z$.

hence by Definition 6.1 either $y_1 \cdot a \neq z$ or $y_2 \cdot a \neq z$. Since neither of these is possible, then $y_1 = y_2 = z$.

Case II. $b \neq z$.

by Theorem 33 there exists an element x such that $a \cdot x = U$

hence $(y_1 \cdot a) \cdot x = (y_2 \cdot a) \cdot x$

by Theorem 31 $y_1 \cdot (a \cdot x) = y_2 \cdot (a \cdot x)$

hence $y_1 \cdot U = y_2 \cdot U$

by Theorem 32.1 $y_1 = y_2$.

THEOREM 36. *The element y of Theorem 34 is uniquely determined by a and b.*

by (A) of Theorem 34 $y = boa$.

Hence the operation o is the inverse of the operation $\cdot$.

THEOREM 37. $(a+b)\cdot c = a\cdot c + b\cdot c$.

Case I. $c = z$.

by Definition 6.2	$(a+b)\cdot c = z,\ a\cdot z = z,\ b\cdot z = z$
by Definitions 1, 2, and 3	$z + z = z$.

Case II. $c \neq z$.

by Postulate 22.2	$(a - b')oc^* = aoc^* - (boc^*)'$
by Definition 6.1	$(a - b')\cdot c = a\cdot c - (b\cdot c)'$
by Definition 3	$(a+b)\cdot c = a\cdot c + b\cdot c$.

THEOREM 38. $a\cdot(b+c) = a\cdot b + a\cdot c$.

by Theorem 37	$(b+c)\cdot a = b\cdot a + c\cdot a$
by Theorem 27	$a\cdot(b+c) = a\cdot b + a\cdot c$.

THEOREM 39. $(aob^*)' = a'ob^*$ *if* $b \neq z$.

by Theorem 26	$zob^* = z$
by Definitions 1 and 2	$(a' - a')ob^* = z$
by Postulate 22.2	$a'ob^* - (aob^*)' = z$
by Definition 1	$a'ob^* - a'ob^* = z$
hence	$a'ob^* - a'ob^* = a'ob^* - (aob^*)'$
by Theorem 25	$a'ob^* = (aob^*)'$.

THEOREM 40. $a - b = (b - a)'$.

by Postulate 13	$(a - b) - (b - a)' = a - [b' - (b - a)']'$
by Postulate 13	$= a - [(b' - b') - a]'$
by Definition 1	$= a - (z - a)'$
	$= a - a'' = a - a = z$ by Definitions 1 and 2 and Postulate 14
by Definition 1	$(a - b) - (a - b) = z$
hence	$(a - b) - (a - b) = (a - b) - (b - a)'$
by Theorem 25	$a - b = (b - a)'$.

THEOREM 41. *If* $b - a = c - a$, *then* $b = c$.

by hypothesis	$b - a = c - a$
by Postulate 11	$(b - a) - a' = (c - a) - a'$
by Postulate 13	$b - (a' - a')' = c - (a' - a')'$
by Definitions 1 and 2	$b - z = c - z$
by RIV	$b = c$.

THEOREM 42. $a - b' = b - a'$.

Case I. $a = b = z$.

by Definition 1 $z - z = z - z = z$.

Case II. $a = z$, $b \neq z$ (also $b = z$, $a \neq z$).

by Definition 2 and Postulate 14 $a - b' = z - b' = b'' = b$

by Definition 2 and RIV $b - a' = b - z = b$.

Case III. $a \neq z$, $b \neq z$, $a - b' = z$.

by hypothesis $a - b' = z$

by Definition 1 $z = b' - b'$

hence $a - b' = b' - b'$

by Theorem 41 $a = b'$

by Postulate 14 $a' = b$

by Definition 1 $b - a' = z$.

Case IV. $a \neq z$, $b \neq z$, $a - b' \neq z$.

by Postulate 20 $a - b' = U \circ (a - b')^*$

by Postulate 22.1 $= U \circ a^* - (U \circ b^*)'$

by Theorem 39 $= U \circ a^* - U' \circ b^*$

by Theorem 39 $= (U' \circ a^*)' - U' \circ b^*$

by Theorem 40 $= [U' \circ b^* - (U' \circ a^*)']'$

by Postulate 22.1 $= [U' \circ (b - a')^*]'$

$= U \circ (b - a')^*$ by Theorem 39 and Definition 5

by Postulate 20 $= b - a'$.

THEOREM 43. $a^{**} = a$.

Case I. $a = z$. For a^{**} to have meaning in this case, a^* must be in K and a^{**} must be in K.

by Postulate 21 $U \circ a^* = a \circ U^*$

hence $U \circ z^* = z \circ U^*$

by Theorem 26 $= z$.

Case II. $a \neq z$.

by Postulate 21 $a \circ a^{**} = a^* \circ a^*$ (A)

by Definition 4 $= U$

by Postulate 22.2 $(a - a^{**}) \circ a^{**} = a \circ a^{**} - (a^{**\prime} \circ a^{**})'$

by Theorem 39 $= a \circ a^{**} - a^{**} \circ a^{**}$

by Definition 4 and (A) $= U - U$

by Definition 1 $= z$ (B)

by Theorem 26 and (B) $[(a-a^{**})oa^{**}]oa^{*}=zoa^{*}=z$
by Postulate 19 $(a-a^{**})o(a^{*}oa^{*})^{*}=z$
by Definition 4 $(a-a^{**})oU^{*}=z$ (E)
now let us suppose $a-a^{**}\neq z$
by Postulate 21, (E) $(a-a^{**})oU^{*}=Uo(a-a^{**})^{*}=z$ (C)
if $a-a^{**}\neq z$, then
by Postulate 18.2 $(a-a^{**})^{*}\neq z$ and $Uo(a-a^{**})^{*}\neq z$
but by (C) we see $Uo(a-a^{**})^{*}=z$
hence the assumption that $a-a^{**}\neq z$ leads to a contradiction and hence

$a-a^{**}=z$
by Theorem 25 $a=a^{**}$.

THEOREM 44. $ao(b-c')^{*}=aob^{*}-(aoc^{*})'$.

Case I. $a=z,\ b\neq z,\ c\neq z,\ b-c'\neq z$.
by Theorem 26 $zo(b-c')^{*}=zob^{*}=zoc^{*}=z$.

Case II. $b-c'=z,\ b\neq z,\ c\neq z,\ a\neq z$. For this case to have meaning $ao(b-c')^{*}$ must be in K.
by Postulate 22.2 $(b-c')oa^{*}=boa^{*}-(coa^{*})'$
by Postulate 21 $ao(b-c')^{*}=aob^{*}-(aoc^{*})'$.

Case III. $b-c'=z,\ b=z,\ c\neq z,\ a\neq z$. The proof of this case is the same as Case II.

Case IV. $b-c'=z,\ b=z,\ c=z,\ a\neq z$. The proof of this case is the same as Case II.

Case V. $b-c'=z,\ b\neq z,\ c\neq z,\ a=z$. The proof of this case is the same as Case II.

Case VI. $a\neq z,\ b\neq z,\ c\neq z,\ b-c'\neq z$.
by Postulate 22.2 $(b-c')oa^{*}=boa^{*}-(coa^{*})'$
by Postulate 21 $ao(b-c')^{*}=aob^{*}-(aoc^{*})'$.

However Theorems 23, 24, 27, 31, 32.1, 32.2, 33, 34, 37 and 38 are the postulates for a field in terms of the operations $+,\cdot$.[4] Hence any system $(K,-,o)$ which satisfies Postulates 11 through 22.2 is a field in terms of the operations $+,\cdot$.

The postulates are examined for independence by exhibiting examples of systems $(K,-,o)$ which fail to satisfy the correspondingly numbered postulates but satisfy the remaining postulates.

---

[4] See reference to Huntington's paper in footnote 3.

Example 11. K is the class of three numbers 0, 1, 2 with $a - b$ and aob satisfying the following multiplication tables. The elements w and u which appear in the tables are elements not in K.

$a-b$	0	1	2
0	0	2	1
1	1	0	w
2	2	w	0

aob	0	1	2
0	u	0	0
1	u	1	2
2	u	2	1.

Example 12. K is the class of all positive rational numbers including zero. $a - b = a + b$ with $z = 0$. $aob = a/b$.

Example 13. K is the class of all positive rational numbers including zero. $a - b = a - b$. $aob = a/b$.

Example 14. K is the class of three numbers 0, 1, 2 with $a - b$ and aob satisfying the following multiplication tables. The element u which appears in the table for o is an element not in K.

$a-b$	0	1	2
0	0	0	0
1	0	0	1
2	0	1	0

aob	0	1	2
0	u	0	0
1	u	1	2
2	u	2	1.

Example 16. K is the class of all integers, positive, negative, and zero. $a - b = a - b$. $aob = a/b$.

Example 17. K is the class of all rational numbers, positive, negative, and zero. $a - b = a - b$. $aob = ab$ with $U = 1$.

Example 18. 1. K is the class of all rational numbers, positive, negative, and zero. $a - b = a - b$. $aob = (a - a)/b$, that is equals zero. Here $U = 0$.

Example 18. 2. K is the class of all rational numbers, positive, negative, and zero. $a - b = a - b$. $aob = a/b$ except $1/a = 0$.

Example 19. K is the class of hypercomplex numbers of the form $\pi 1 + \omega i + \rho j$ where π, ω, ρ are rational numbers, positive, negative, and zero. $a - b = a - b$.

$$aob = \frac{1}{\pi_2{}^2 + \rho_2{}^2 + (\omega_2 - \rho_2)^2} (\pi_1 1 + \omega_1 i + \rho_1 j)(\pi_2 1 + \omega_2 i + \rho_2 j),$$

where the product of the coefficients shall be the ordinary product of rational numbers and the "units" shall follow the table,

	1	i	j
1	1	$-i$	$-j$
i	i	1	-1
j	j	-1	2.

Example 21. K is the class of quaternions. $a - b = a - b$.

$$aob = \frac{1}{\pi_2^2 + \omega_2^2 + \rho_2^2 + \delta_2^2} (\alpha_2\delta + \beta_2\gamma + i\omega_2 + j\rho_2 + k\pi_2)(\alpha_1\delta + \beta_1\gamma + i\omega_1 + j\rho_1 + k\pi_1)$$

where the product of the coefficients shall be the ordinary product of rational numbers, while the "units" shall obey the table

	1	i	j	k
1	1	$-i$	$-j$	$-k$
i	i	1	$-k$	j
j	j	k	1	$-i$
k	k	$-j$	i	1.

Example 22.2. K is the class of two numbers 0, 1. $a - b$ and aob satisfy the following tables. The element u in the table for o is an element not in K.

$a - b$	0	1
0	0	1
1	1	0

aob	0	1
0	u	1
1	u	1.

Example E. K is the class consisting of 0 only with $0 - 0 = 0$ and 0o0 undefined.

HARVARD UNIVERSITY.